

10th Edition

Automotive Handbook



BOSCH

Invented for life

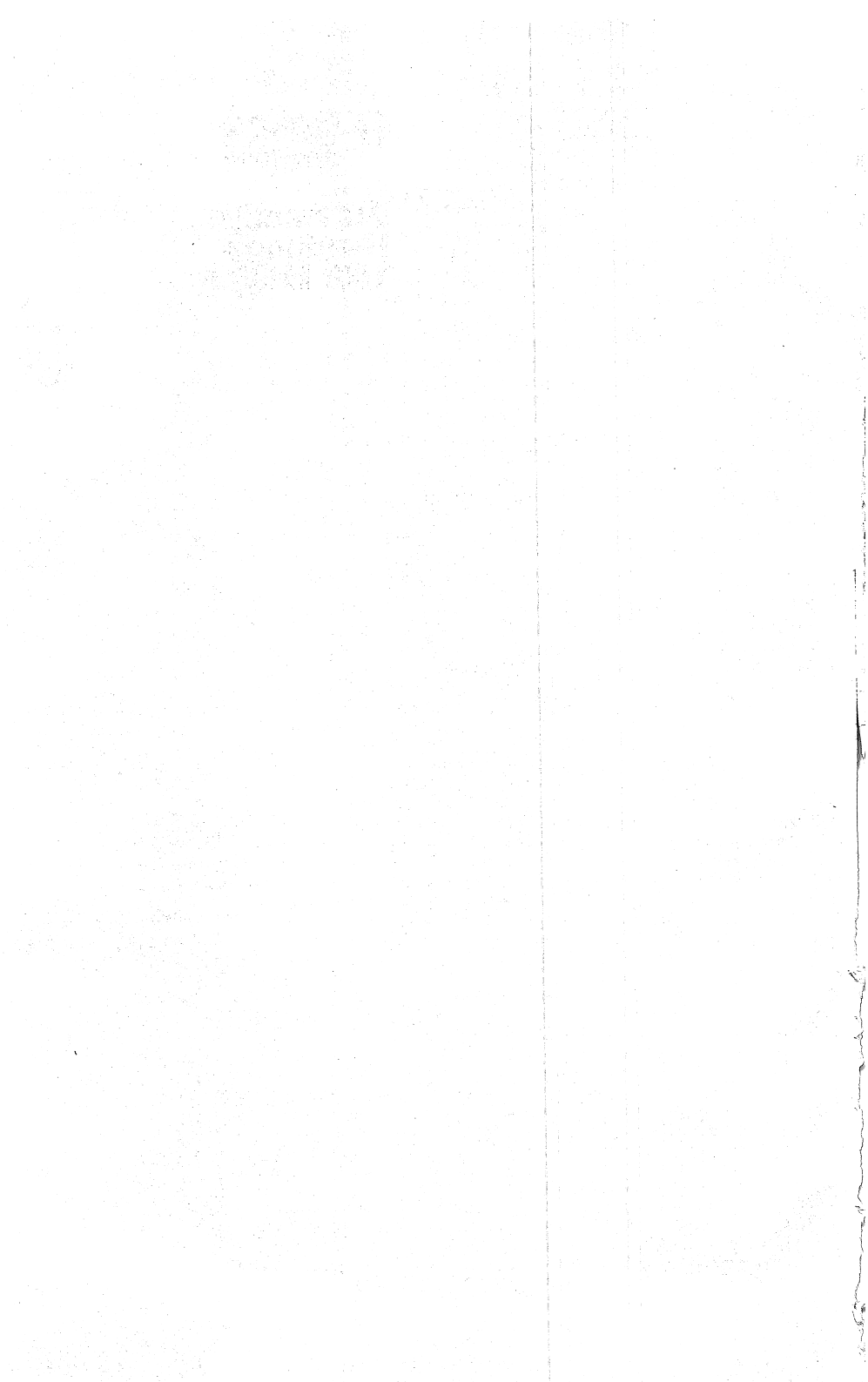
WILEY



BOSCH

Invented for life

Automotive Handbook 10th Edition





BOSCH

Invented for life

Automotive Handbook

Imprint

Published by:

© Robert Bosch GmbH, 2018
Postfach 410960,
D-76225 Karlsruhe.
Business Division
Automotive Aftermarket.

Scientific advisor and editor

Prof. Dr.-Ing. Konrad Reif,
Duale Hochschule Baden-Württemberg,
Ravensburg, Campus Friedrichshafen,
Academic Program Director
Automotive Electronics and
Mechatronic Systems.

Editorial staff

Dipl.-Ing. Karl-Heinz Dietsche;
Technischer Redakteur (FH)
Matthias Thiess.

Type setting

Dipl.-Ing. Karl-Heinz Dietsche.

Translation

STAR Deutschland GmbH
Member of STAR Group.

10th Edition, revised and extended,
September 2018.

All rights reserved.

Printed in Germany.
Imprimé en Allemagne

Distribution

John Wiley & Sons Ltd,
The Atrium,
Southern Gate, Chichester,
West Sussex PO19 8SQ, England
Telephone (+44) 1243 779777
Internet: www.wiley.com or
www.wileyeurope.com
Email: cs-books@wiley.co.uk

ISBN 978-1-119-53081-7

Bibliographic information published
by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists
this publication in the Deutsche Nationalbibliografie; detailed bibliographic
data are available in the Internet at
<http://dnb.d-nb.de>.

Authors

Dipl.-Ing. Karl-Heinz Dietsche,
Prof. Dr.-Ing. Konrad Reif and
approx. 200 authors from industry
and the university and college sector.

Technical graphics

Bauer & Partner, Gesellschaft für
technische Grafik mbH, Stuttgart;
Schwarz Technische Grafik, Leonberg.

Reproduction, duplication and translation of this publication, including excerpts therefrom, is only to ensue with our prior written consent and with particulars of source. Figures, descriptions, schematic diagrams, and other data are for explanatory purposes and illustration of the text only. They cannot be used as the basis for the design, installation, or specification of products.

The comments reflect solely the personal views of the respective authors. Comments relating to legal standards make no claims of being complete and correct. The contents of the Handbook do not make any statement as to whether vehicles and vehicle component parts are approvable in an individual case and whether they comply with the applicable legal requirements in each case.
Subject to alteration and amendment.

The brand names given in the contents are only examples and do not represent the classification or preference for a particular manufacturer. Trademarks are not identified as such.

The following companies kindly placed text and picture matter and diagrams at our disposal:

Automotive Lighting Reutlingen GmbH.

BASF Ludwigshafen.

Bosch Engineering GmbH, Abstatt.

Brose Schließsysteme GmbH & Co. KG, Wuppertal.

Daimler AG, Stuttgart.

Daimler AG, Sindelfingen.

Dr. Ing. h.c. F. Porsche AG, Weissach.

Eberspächer Climate Control Systems GmbH & Co. KG, Esslingen.

ETAS GmbH, Stuttgart.

Gates GmbH, Aachen.

IAV GmbH, Berlin.

iwis motorsysteme GmbH & Co. KG, München.

J. Eberspächer GmbH & Co. KG, Esslingen.

Johnson Controls Autobatterie GmbH & Co. KGaA, Hannover.

Knorr-Bremse Systeme für Nutzfahrzeuge GmbH, Schwieberdingen.

MAHLE Behr GmbH & Co. KG, Stuttgart.

MAN Nutzfahrzeuge Gruppe.

MANN+HUMMEL GmbH, Ludwigsburg.

Michelin Reifenwerke AG & Co. KGaA, Karlsruhe.

Robert Bosch Automotive Steering GmbH Schwäbisch Gmünd.

Robert Bosch Battery Systems GmbH, Stuttgart.

Robert Bosch Starter Motors Generators GmbH, Schwieberdingen.

Saint-Gobain Sekurit International, Herzogenrath.

Trelleborg Sealing Solutions Germany GmbH, Stuttgart.

Volkswagen AG, Wolfsburg.

ZF Friedrichshafen AG, Friedrichshafen.

Duale Hochschule Baden-Württemberg, Ravensburg, Campus Friedrichshafen.

fka Forschungsgesellschaft Kraftfahrzeugwesen mbH Aachen.

Friedrich-Alexander-Universität Erlangen-Nürnberg.

Hochschule Esslingen.

Hochschule Karlsruhe – Technik und Wirtschaft.

Hochschule München.

Karlsruher Institut für Technologie (KIT).

Reinhold-Würth-Hochschule, Künzelsau.

Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen.

Foreword to the 10th Edition in English

Automotive engineering has become an extremely complex field over the last few years and decades. It is becoming increasingly more difficult to command an overview of the entire field and to maintain constant access to the subjects which are significant to automotive engineering. Many of these new subjects have in the meantime been covered in great detail in the wealth of available specialist literature. However, for those readers who wish to approach one of these new subjects for the first time, the available literature is neither easily manageable nor is it readable within the available timeframe.

This is where the Automotive Handbook comes in useful. It is structured in such a way as to be easily accessible even to those readers who are new to any individual subject. The most important subjects relevant to automotive engineering have been compiled in a compact, easily understandable, and practically relevant form. This is because all the content has been written by experts at Bosch and other vehicle manufacturers and suppliers who work in the very fields covered in the Handbook.

In this the 10th Edition in English many subjects have been included for the first time, completely revised or enhanced. Included for the first time among others are subjects pertaining to hybrid drives, while the drivetrain has been completely revised and significantly enhanced. The Handbook has however also been revised, expanded and updated in many other places to make the contents even more understandable to the reader. More than 200 pages have been added to this latest edition. In spite of the large number of contributing authors, every effort has been made to provide a uniform presentation and to maintain consistent classifications and nomenclature within its pages.

This individual chapters of this publication have been laid out gradually on receipt of the authors' manuscripts. Care has been taken to ensure that the graphics where possible are positioned in such a way that they appear with the accompanying text on the same page or the facing page and consequently both are visible to the reader in a single glance. Shifting the pagination at a later stage by one page would undo this text/graphic reference. This does however mean that there are blank pages in some places.

This latest edition could not have been completed without the outstanding support of many individuals. Firstly, we would like to thank the authors of the individual articles, who, with great care and patience, have succeeded in delivering on schedule chapters of the highest substance and quality. Finally, we would like to thank all those readers who have provided useful suggestions and advice on corrections.

Friedrichshafen and Karlsruhe, September 2018,

Scientific advisory board and
editorship.

Contents

Basic principles			
Quantities and units	24	Wave propagation	108
SI units	24	Electric effects in metallic	
Legal units	26	conductors	111
Further units	33	Electronics	118
Natural constants	36	Basic principles of semiconductor	
Mathematical signs and symbols	37	technology	118
Greek alphabet	37	Discrete semiconductor	
Basic principles of mechanics	38	devices	121
Rectilinear and rotary motion	38	Monolithic integrated circuits	137
Statics	41	Chemistry	138
Calculation of strength	43	Elements	138
Friction	45	Chemical bonds	144
Vibrations and oscillations	48	Substances	148
Terms	48	Substance concentrations	150
Equations	50	Reaction of substances	151
Vibration reduction	51	Electrochemistry	158
Modal analysis	52	Electrolytic conduction and	
Acoustics	54	electrolysis	158
General terminology	54	Applications	159
Measured quantities for noise			
emissions	57	Mathematics and methods	
Measured quantities for noise		Mathematics	164
immissions	58	Numbers	164
Perceived sound levels	59	Functions	164
Optical technology	62	Equations in the plane triangle	170
Electromagnetic radiation	62	Complex numbers	170
Geometric optics	62	Coordinate systems	171
Wave optics	65	Vectors	173
Lighting quantities	67	Differential and integral calculus	175
Laser technology	70	Linear differential equations	178
Optical fibers/waveguides	71	Laplace transform	179
Fluid mechanics	74	Fourier transform	180
Hydrostatics	74	Matrix calculus	182
Flow mechanics	75	Technical statistics	186
Thermodynamics	78	Descriptive statistics	186
Basic principles	78	Probability calculus	188
Laws of thermodynamics	80	Characteristic value in quality	
Equations of state	81	assurance and in quality	
Changes of state	83	management	192
Cycles and technical applications	83	Finite-element method	194
Heat transfer	88	Applications	194
Electrical engineering	94	FEM examples	197
Electromagnetic fields	94	Control engineering	202
Electric field	94	Terms and definitions	202
Direct current and direct voltage	96	Control-engineering transfer	
Time-dependent current	99	elements	203
Magnetic field	100	Designing a control task	204
Magnetic field and electric		Adaptive controllers	206
current	104		

Materials		Machine parts	
Materials	208	Tolerances	352
Material parameters	208	Form deviations	352
Material groups	220	Geometrical Product Specifications	352
Metallic materials	221	Axles and shafts	357
EN metallurgy standards	236	Function and application	357
Magnetic materials	240	Dimensioning	357
Nonmetallic inorganic materials	253	Design	359
Composite materials	253	Springs	360
Plastics	256	Basic principles	360
Heat treatment of metallic materials	278	Metal springs	362
Hardness	278	Friction bearings	368
Heat-treatment processes	281	Features	368
Thermochemical treatment	287	Hydrostatic friction bearings	368
Corrosion and corrosion protection	290	Hydrodynamic friction bearings	369
Corrosion processes	290	Self-lubricating bearings made of metal	372
Types of corrosion	291	Self-lubricating bearings made of plastic	374
Corrosion testing	292	Rolling bearings	376
Corrosion protection	294	Application	376
Deposits and coatings	298	General principles	376
Deposits	298	Selection of rolling bearings	377
Coatings	303	Calculation of load capacity	380
Operating fluids		Seals	382
Lubricants	304	Seal technology	382
Terms and definitions	304	O-ring	384
Engine oils	310	Molded seal	391
Transmission oils	314	Flat seals	393
Lubricating oils	315	Sealing caps	394
Lubricating greases	315	Radial shaft seal	396
Coolants for internal-combustion engines	319	Pneumatic seals	404
Requirements	319	Two-component seals	406
Composition of coolants	319	Rod and piston sealing systems	409
Types of radiator protection agents	320	Elastomer materials	413
Brake fluids	322	Joining and bonding techniques	
Requirements	322	Detachable connections	424
Characteristics	323	Positive or form-closed joints	424
Chemical composition	324	Frictional joints	428
Fuels	326	Threaded fasteners	435
Overview	326	Snap-on connections on plastic components	445
Characteristics	327	Permanent joints	448
Gasoline (petrol)	328	Welding	448
Diesel fuel	333	Soldering	452
Natural gas	341	Adhesive technologies	453
Liquefied petroleum gas	342	Riveting	455
Hydrogen	343	Penetration-clinching processes	457
e-fuels	343		
Ethers	344		
Urea/water solution	348		
Refrigerants for air-conditioning systems	349		
Development up to today	349		
Alternative possibilities	350		

Internal-combustion engines

Internal-combustion engines	460
Thermal engines	460
Mixture formation, combustion, emissions	464
Gasoline engine	464
Diesel engine	475
Mixed forms and alternative operating strategies	480
Charge cycle and supercharging	482
Charge cycle	482
Variable valve timing	485
Supercharging processes	489
Exhaust-gas recirculation	492
Reciprocating-piston engine	494
Components	494
Reciprocating-piston engine types	508
Crankshaft-assembly design	509
Tribology and friction	517
Empirical values and data for calculation	521
Exhaust emissions	534
Historic measures	534
Constituents of engine exhaust gas	534
Basic principles of catalytic exhaust-gas treatment	536
Catalytic processes for exhaust-gas treatment	537
Belt drives	538
Friction belt drives	538
Positive belt drives	543
Chain drives	548
Overview	548
Timing chains	548
Sprockets	552
Chain-tensioning and chain-guide elements	552
Engine cooling	554
Air cooling	554
Water cooling	555
Intercooling (charge-air cooling)	561
Exhaust-gas cooling	562
Oil and fuel cooling	563
Modularization	565
Intelligent thermal management	567
Engine lubrication	568
Forced-feed lubrication system	568
Components	569
Thermal management	571
Air-intake systems and intake manifolds	572
Overview	572
Passenger-car air-intake system	573

Passenger-car intake manifolds	581
Motorcycle intake manifolds	586
Commercial-vehicle air-intake system	586
Commercial-vehicle intake manifolds	589
Turbochargers and superchargers	590
Superchargers (mechanically driven)	590
Pressure-wave superchargers	592
Exhaust-gas turbochargers	593
Complex supercharging systems	601
Exhaust-gas system	606
Purpose and design	606
Exhaust manifold	608
Catalytic converter	608
Particulate filter	609
Mufflers	610
Connecting elements	612
Acoustic tuning devices	613
Commercial-vehicle exhaust-gas systems	614

Management for spark-ignition engines

Management for spark-ignition engines	616
Description of the engine management system	616
System overview	617
Motronic versions for passenger cars	621
Motronic versions for motorcycles	625
Cylinder charge	628
Component parts	628
Control of cylinder charge	629
Recording of air charge	631
Cylinder shutoff	633
Fuel supply	636
Fuel supply and delivery with intake-manifold injection	636
Fuel supply and delivery with gasoline direct injection	637
Evaporative-emissions control system	639
Gasoline filter	641
Electric fuel pump	643
High-pressure pumps for gasoline direct injection	646
Fuel rail	649
Fuel-pressure regulator	651
Fuel-pressure attenuator	652

Mixture formation	654	Time-controlled single-cylinder pump systems	754
Basic principles	654	Unit-injector system for passenger cars	754
Mixture-formation systems	656	Unit-injector system for commercial vehicles	756
Intake-manifold injection	657	Unit-pump system for commercial vehicles	757
Gasoline direct injection	659	Electronic control	757
Fuel injectors	662	Diesel distributor injection pumps	758
Ignition	670	Axial-piston distributor pumps	758
Basic principles	670	Radial-piston distributor pumps	761
Moment of ignition	671	Fuel-injection system	763
Ignition systems	676	Start-assist systems for diesel engines	764
Ignition coil	678	Preheating systems for passenger cars and light commercial vehicles	764
Spark plug	681	Flame start systems for commercial-vehicle diesel engines	768
Catalytic exhaust-gas treatment	690	Grid heater for commercial-vehicle diesel engines	771
Catalytic converter	690	Exhaust-gas treatment	772
λ -control	695	Catalytic oxidation	772
Particulate filter for gasoline engines	697	Filtration of particulates	773
Motor sport	698	Catalytic reduction of nitrous fumes	775
Requirements	698	On-Board Diagnostics (OBD)	779
Components	698	Drivetrain	
ECUs	700	Conventional drivetrain	780
Pure motor-sport developments	700	Drivetrain elements	780
Alternative gasoline-engine operation		Function	780
LPG operation	702	Arrangement of engine	783
Applications	702	Vibration isolation	784
Design	703	Power take-up elements	786
LPG systems	705	Single clutch	786
Components	707	Dual clutch	787
Engines fueled by natural gas	710	Dry and wet clutches	788
Applications	710	Hydrodynamic torque converter	788
Design	712	Transmission	791
Components	713	Step-up gear units	791
Engines fueled by alcohol	716	Gear-shift elements	793
Application	716	Manually shifted transmissions	795
Ethanol as an alternative fuel	717	Automated manual transmissions	797
Flex-fuel concepts for the markets	719	Dual-clutch transmissions	798
Flex-fuel components	720	Automatic planetary transmissions	799
Management for diesel engines		Continuously variable transmissions	800
Management for diesel engines	722	Commercial-vehicle transmissions	802
Description of the engine-management system	722	Electronic transmission control	804
Electronic Diesel Control	723	Differentials	808
Low-pressure fuel supply	726	Final-drive differentials	808
Fuel supply and delivery	726		
Diesel filter	732		
Common-rail injection system	734		
System overview	734		
Injectors	739		
High-pressure pumps	747		
Rail	753		

Drivetrain with hybrid drives	812	Emission legislation for motorcycles	894
Features	812	Exhaust-gas measuring techniques	900
Functionalities of hybrid drives	812	Exhaust-gas testing	900
Classification based on proportion of electric power	814	Exhaust-gas measuring devices	903
Classification based on drive structure	815	Diesel smoke-emission test	907
Drivetrain with electric drives	820	Evaporative-emissions test	909
Pure battery-operated electric drive	820	Diagnostics	912
Electric drive with range extender	821	Monitoring in driving mode	912
		On-Board Diagnostics	914
		OBD functions	920
		OBD requirements for heavy commercial vehicles	923
		On-Board Diagnostics for motorcycles	927
		ECU diagnostics and Service Information System	928
Alternative drives			
Hybrid drives	822	Vehicle physics	
Features	822	Basic terms of automotive engineering	930
Drive structures	823	Basic terms of vehicle handling	930
Control of hybrid vehicles	829	Motor-vehicle dynamics	940
Regenerative braking system	832	Dynamics of linear motion	940
Boost-recuperation system 48 V	834	Adhesion to road surface	946
Motivation	834	Accelerating and braking	947
Requirements imposed on a 48 V electric machine	836	Actions: Reaction, braking and stopping	948
Choice of concept for the electric machine	838	Passing (overtaking)	950
Thermal situation and rating	840	Fuel consumption	952
Machine performance in the system network	841	Dynamics of lateral motion	958
Overall view of the 48 V system	842	Special operating dynamics for commercial vehicles	964
Electric machines for motor-vehicle propulsion	844	Driving-dynamics test procedures as per ISO	967
Requirements	844	Vehicle aerodynamics	974
Types of electric machines	847	Aerodynamic forces	974
Influencing factors on operational performance	849	Tasks of vehicle aerodynamics	977
Losses in electric machines	854	Vehicle wind tunnels	982
Cooling of the electric machine	856	Vehicle acoustics	988
Fuel cells for the vehicle drive	858	Legal requirements	988
Functioning principle	858	Development work on vehicle acoustics	990
Functioning principle of the fuel-cell system	860	Sources of noise	991
Functioning principle of the drivetrain	862	Sound design	993
Emission-control and diagnosis legislation		Chassis systems	
Emission-control legislation	864	Chassis systems	996
Overview	864	Definition and design	996
Emission-control legislation for passenger cars and light commercial vehicles	866	Basic principles	997
Test cycles for passenger cars and light commercial vehicles	880	Suspension	1006
Emission-control legislation for heavy commercial vehicles	885	Basic principles	1006
Test cycles for heavy commercial vehicles	890	Types of spring	1009
		Suspension systems	1014

Shock absorbers and vibration absorbers	1018	Brake systems for commercial vehicles	1106
Shock absorbers	1018	System overview	1106
Vibration absorbers	1025	Components of commercial-vehicle brake systems	1109
Wheel suspensions	1026	Electronically controlled brake system	1118
Basic principles	1026	Continuous-operation brake systems	1122
Kinematics and elastokinematics	1026	Wheel brakes	1126
Basic categories of wheel suspensions	1028	Disk brakes for passenger cars	1126
Wheels	1034	Disk brakes for commercial vehicles	1131
Function and requirements	1034	Drum brakes	1132
Structure	1034		
Design criteria	1039	Chassis control and active safety	
Designation for passenger-car wheels	1039	Wheel-slip control systems	1136
Materials for wheels	1040	Function and requirements	1136
Manufacturing processes	1042	Control systems	1137
Wheel design variations	1045	ABS/TCS systems for passenger cars	1141
Stress and testing of wheels	1047	ABS/TCS systems for commercial vehicles	1145
Tires	1050	Driving-dynamics control system	1148
Function and requirements	1050	Function	1148
Tire construction	1051	Requirements	1149
Tire inflation pressure	1053	Operating principle	1150
Tire tread	1054	Typical driving maneuver	1150
Force transmission	1056	Structure of the overall system	1151
Tire grip	1957	System components	1160
Rolling resistance	1060	Special driving-dynamics control system for commercial vehicles	1163
Tire designation	1062	Automatic brake-system operations	1168
EU tire label	1065	Integrated driving-dynamics control systems	1172
Winter tires	1066	Overview	1172
Development aims	1067	Functions	1172
Tire tests	1067	System architecture	1175
Tire-pressure monitoring systems	1068		
Rotary seal for tire-pressure control	1070	Passive safety	
Steering	1072	Passive safety	1178
Purpose of automotive steering systems	1072	Phases of vehicle safety	1178
Steering-system requirements	1073	System architecture, passive safety	1180
Types of steering box	1075	Legal regulations and consumer tests	1181
Power-assisted steering systems for passenger cars	1076	Occupant-protection systems	1182
Power-assisted steering systems for commercial vehicles	1083	Functions	1182
Brake systems	1086	Retention systems and actuators	1184
Definitions and principles	1086	Occupant sensing	1188
Legal regulations	1091	Integral-safety systems	1190
Structure and organization of brake systems	1094	Overview	1190
Brake systems for passenger cars and light utility vehicles	1096	Solution approaches	1191
Subdivision of passenger-car brake systems	1096	Outlook	1195
Components of the passenger-car brake system	1097		
Electrohydraulic brake	1104		

Vehicle bodies

Vehicle bodies, passenger cars	1196
Main dimensions	1196
Body design	1198
Body	1199
Safety	1203
Vehicle bodies, commercial vehicles	1208
Classification of commercial vehicles	1208
Light utility vans	1209
Medium- and heavy-duty trucks and tractor units	1209
Buses	1214
Active and passive safety in commercial vehicles	1216
Lighting equipment	1218
Functions	1218
Regulations and equipment	1219
Light sources	1220
Main-light functions	1221
Main-light functions for Europe	1224
Main-light functions for the USA	1227
Definitions and terms	1229
Technical design variations of headlamps	1230
Installation and regulations for signal lamps	1243
Technical design variations for lamps	1247
Headlamp leveling control	1251
Headlamp adjustment	1252
Automotive glazing	1258
The material properties of glass	1258
Automotive windshield and window glass	1259
Functional design glazing	1260
Windshield, rear-window and headlamp cleaning systems	1264
Windshield wiper systems	1264
Rear-window wiper systems	1270
Windshield and rear-window washer systems	1271
Headlamp washer systems	1271

Vehicle security systems

Locking systems	1272
Task and design	1272
Access authorization	1273
Lock design	1274
Requirements	1276
Lock layout	1276
Safety functions	1280
Development history	1281

Anti-theft systems	1284
Electronic vehicle immobilizer	1284
Theft-alarm system	1288

Automotive electrics

Conventional vehicle electrical systems	1290
Electrical power supply	1290
Electrical-system structures	1296
Electrical-system parameters	1297
Electrical energy management	1300
Vehicle electrical systems for hybrid and electric vehicles	1304
Vehicle electrical systems for mild and full hybrid vehicles	1304
Vehicle electrical systems for plug-in hybrid and electric vehicles	1307
Charging strategy	1308
Starter batteries	1310
Requirements	1310
Battery design	1312
Charging and discharging	1314
Battery characteristics	1315
Battery types	1316
Using the battery	1319
Drive batteries	1322
Requirements	1322
Storage technologies	1323
Basic structure of a battery system	1324
Components of a lithium-ion battery system	1325
Thermal management	1329
Electric machines	1330
Energy conversion	1330
Systematics of rotating electric machines	1331
Definition of geometric quantities	1331
Direct-current machines	1332
Asynchronous machine	1338
Synchronous machine	1340
Switched reluctance motor	1345
Electronic motors	1348
Energy-efficiency classes	1349
Three-phase current system	1350
Alternators	1354
Electric power generation	1354
Voltage regulation	1359
Characteristic alternator values	1361
Operating conditions	1362
Efficiency	1363
Alternator versions	1364

Starting systems	1368	LIN	1466
Starter	1368	Ethernet	1469
Triggering the starter	1375	PSI5	1471
Start/stop system	1376	MOST	1473
Wiring harnesses and plug-in connections	1378	Serial Wire Ring	1477
Wiring harnesses	1378	Architecture of electronic systems	1482
Plug-in connections	1380	General	1482
Electromagnetic compatibility	1384	Development of E/E architectures	1490
Requirements	1384	Summary and outlook	1496
Interference emission and interference immunity	1385	Automotive sensors	1498
EMC-oriented development	1388	Basic principles	1498
EMC measuring techniques	1389	Position and angular-position sensors	1504
EMC test methods	1390	Rpm sensors	1518
EMC simulation	1393	Oscillation gyrometers	1523
Legal requirements and standards	1394	Flowmeters	1526
Symbols and circuit diagrams	1398	Acceleration and vibration sensors	1530
Circuit symbols	1398	Pressure sensors	1536
Circuit diagrams	1405	Temperature sensors	1540
Terminal designations	1417	Torque sensor	1544
Automotive electronics		Force sensor	1545
Electronic control unit (ECU)	1420	Gas and concentration sensors	1546
Functions	1420	Optoelectronic sensors	1556
Requirements	1420	Ultrasonic sensor	1559
ECU components	1421	Radar sensors	1560
Data processing	1427	Lidar sensors	1564
Automotive software engineering	1430	Video sensors	1567
Motivation	1430	Mechatronics	1572
Design of software in motor vehicles	1431	Mechatronic systems and components	1572
Important standards for software in motor vehicles	1431	Development methodology	1574
The development process	1434	Outlook	1577
Quality assurance in software development	1438	Comfort and convenience	
Workflows of software development in motor vehicles	1438	Passenger-compartment climate control	1578
Modeling and simulation of software functions	1440	Climate-control requirements	1578
Design and implementation of software functions	1443	Design and operating principle of the A/C unit	1578
Integration and testing of software and ECUs	1444	Climate-control systems	1581
Calibration of software functions	1446	Climate control for hybrid and electric vehicles	1582
Outlook	1448	Cabin filters	1583
Automotive networking	1450	Auxiliary heater systems	1584
Bus systems	1450	Comfort and convenience systems	
Technical principles	1451	in the door and roof areas	1590
Buses in motor vehicles	1458	Power-window systems	1590
CAN	1458	Sunroof systems	1592
FlexRay	1463	Comfort and convenience functions	
		in the passenger compartment	1594
		Electrical seat adjustment	1594
		Electrical steering-wheel adjustment	1595

User interfaces, telematics and multimedia

Display and control	1596
Interaction channels	1596
Instrumentation	1597
Display types	1601
Radio and TV reception in motor vehicles	1606
Wireless signal transmission	1606
Radio tuners	1609
Traffic telematics	1613
Transmission paths	1613
Standardization	1613
Information recording	1614

Driver-assistance systems

Driver-assistance systems	1616
Introduction – driver assistance	1616
Computer vision	1624
Camera model	1624
Image processing	1627
Vehicle navigation	1632
Navigation systems	1632
Functions of navigation	1632
Digital map	1635
Night-vision systems	1636
Applications	1636
Far-infrared systems	1636
Near-infrared systems	1637
Parking and maneuvering systems	1642
Application	1642
Ultrasonic parking aid	1642
Ultrasonic parking assistant	1645
Video systems	1648
Adaptive Cruise Control	1652
Function	1652
Design and function	1652
Control algorithms	1654
Area of application and functional expansions	1655
Current ongoing developments	1657
Lane assistance	1658
Lane departure warning	1658
Lane-keeping support	1659
Roadworks assistant	1660
Bottleneck/constriction assistant	1660
Lane-change assistant	1662
Blind-spot assistant	1662
Lane-change assistant	1663
Further developments	1664
Emergency-braking systems	1665
Emergency brake assist and automatic emergency braking	1665

Intersection assistant	1668
Motivation	1668
Intersection-assistance systems	1669
Active pedestrian protection	1671
High-beam assistant	1673
Future of automated driving	1674
Long-term goal – autonomous driving	1674
Obstacles to automated driving	1675
Stages on the road to autonomous driving	1676

Appendices

Technical terms	1680
Abbreviations	1766

Authors

Unless otherwise stated, the authors are employees of Bosch.

Basic principles

Quantities and units

Prof. Dr. rer. nat. Susanne Schandl,
Duale Hochschule Baden-Württemberg,
Ravensburg, Campus Friedrichshafen.

Basic principles of mechanics

Prof. Dr.-Ing. Horst Haberhauer,
Hochschule Esslingen.

Vibrations and oscillations

Dipl.-Ing. Sebastian Loos,
Rheinisch-Westfälische Technische
Hochschule (RWTH) Aachen.

Acoustics

Dipl.-Ing. Hans-Martin Gerhard,
Dr. Ing. h.c. F. Porsche AG, Weissach.

Optical technology

Dipl.-Phys. Stefanie Mayer.

Fluid mechanics

Prof. Dr.-Ing. Horst Haberhauer,
Hochschule Esslingen.

Thermodynamics

Dr.-Ing. Ingo Stotz.

Electrical engineering

Dr.-Ing. Hans Roßmanith,
Friedrich-Alexander-Universität
Erlangen-Nuremberg;
Prof. Dr.-Ing. Klemens Gintner,
Hochschule Karlsruhe –
Technik und Wirtschaft.

Electronics

Prof. Dr.-Ing. Klemens Gintner,
Hochschule Karlsruhe –
Technik und Wirtschaft;
Dr. rer. nat. Ulrich Schaefer.

Chemistry

Dr. rer. nat. Jörg Ullmann.

Electrochemistry

Prof. Dr.-Ing. Matthias E. Rebhan,
Hochschule München.

Mathematics and methods

Mathematics, Technical statistics

Prof. Dr.-Ing. Matthias E. Rebhan,
Hochschule München.

Finite-element method

Prof. Dipl.-Ing. Peter Groth,
Hochschule Esslingen.

Control engineering

Dr.-Ing. Wolf-Dieter Gruhle,
ZF Friedrichshafen AG,
Friedrichshafen.

Materials

Materials

Dr.-Ing. Hagen Kuckert;
Dr. rer. nat. Jörg Ullmann;
Dr. rer. nat. Witold Pieper;
Dr. rer. nat. Waldemar Draxler;
Dipl.-Ing. Angelika Schubert;
Dipl.-Ing. Gert Lindemann;
Dr.-Ing. Carsten Tüchert;
Dr.-Ing. Sven Robert Raisch;
Dr.-Ing. Reiner Lützelzer;
Dr. rer. nat. Jörg Bettenhausen;
Dr.-Ing. Gerrit Hülde;
Dipl.-Ing. Cornelius Gaida;
Dipl.-Phys. Klaus-Volker Schütt.

Heat treatment of metallic materials

Dr.-Ing. Thomas Waldenmaier;
Dr.-Ing. Jochen Schwarzer.

Corrosion and corrosion protection

Dipl.-Ing. (FH) Thomas Jäger.

Deposits and coatings

Dr. Helmut Schmidt;
Dipl.-Ing. (FH) Hellmut Schmid;
Dipl.-Ing. (FH) Susanne Lucas.

Operating fluids

Lubricants

Dr. rer. nat. Gerd Dornhöfer.

Coolants for internal-combustion engines

Dr. rer. nat. Oliver Mamber,
MAHLE Behr GmbH & Co. KG,
Stuttgart.

Brake fluids

Dr. rer. nat. Michael Hilden;
Dr. rer. nat. Harald A. Dietl,
BASF Ludwigshafen.

Fuels

Dr. rer. nat. Jörg Ullmann.

Urea/water solution

Thilo Schulz, M. Sc.;
Dipl.-Ing. (FH) Manfred Fritz.

Refrigerants for air-conditioning systems

Prof. Dr.-Ing. Stephan Engelking,
Duale Hochschule Baden-Württemberg,
Ravensburg, Campus Friedrichshafen.

Machine parts

Tolerances, axles and shafts, springs, friction bearings

Prof. Dr.-Ing. Horst Haberhauer,
Hochschule Esslingen.

Roller bearings

Dr.-Ing. Zhenhuan Wu.

Seals

Dipl.-Ing. (FH) Ulrich Schmid;
Dipl.-Ing. Sascha Bürkle;
Thorsten Nitzte, M. Eng.;
Claudius Göller, MBE (DHBW);
Dipl.-Ing. Wolfgang Böhrer;
Frank Kleemann;
Sergio Amorim, M. Sc.;
Felix Schädler.
All authoris of the chapter Seals
from Trelleborg Sealing Solutions
Germany GmbH, Stuttgart.

Joining and bonding techniques

Detachable connections

Prof. Dr.-Ing. Horst Haberhauer,
Hochschule Esslingen;
Dipl.-Ing. Rolf Bald;
Dr.-Ing. Sven Robert Raisch.

Permanent joints

Dr.-Ing. Knud Nörenberg,
Volkswagen AG, Wolfsburg;
Dr. rer. nat. Patrick Stihler.

Internal-combustion engines

*Internal-combustion engines,
mixture formation,
charge cycle and supercharging,
reciprocating-piston engine*
Prof. Dr. sc. techn. Thomas Koch,
Karlsruher Institut für Technologie (KIT);
Tobias Michler, M. Sc.,
Karlsruher Institut für Technologie (KIT);
Dipl.-Ing. Heijo Oelschlegel,
Daimler AG, Stuttgart;
Dr.-Ing. Otmar Scharrer,
MAHLE Behr GmbH & Co. KG, Stuttgart.

Exhaust-gas emissions

Dr. rer. nat. Christoph Osemann;
Dr.-Ing. Hartmut Lüders.

Belt drives

Dipl.-Ing. (FH) Mario Backhaus,
Gates GmbH, Aachen.

Chain drives

Dr.-Ing. Thomas Fink,
iwis motorsysteme GmbH & Co. KG,
Munich;
Dr.-Ing. Peter Bauer,
iwis motorsysteme GmbH & Co. KG,
Munich.

Engine cooling

Dipl.-Ing. (FH) Ralf-Holger Schink,
MAHLE Behr GmbH & Co. KG, Stuttgart;
Dr.-Ing. Otmar Scharrer,
MAHLE Behr GmbH & Co. KG, Stuttgart.

Engine lubrication

Dipl.-Ing. Markus Kolczyk,
MANN+HUMMEL GmbH, Ludwigsburg;
Dr.-Ing. Harald Banzhaf,
MANN+HUMMEL GmbH, Ludwigsburg.

Air-intake systems and intake manifolds

Dipl.-Ing. Andreas Weber,
MANN+HUMMEL GmbH, Ludwigsburg;
Dipl.-Ing. Andreas Pelz,
MANN+HUMMEL GmbH, Ludwigsburg;
Dipl.-Ing. Markus Kolczyk,
MANN+HUMMEL GmbH, Ludwigsburg;
Dipl. Ing. (FH) Alexander Korn,
MANN+HUMMEL GmbH, Ludwigsburg;
Dipl. Ing. (FH) Matthias Alex,
MANN+HUMMEL GmbH, Ludwigsburg;
Dr.-Ing. Matthias Teschner,
MANN+HUMMEL GmbH, Ludwigsburg;
Dipl. Ing. (FH) Mario Rieger,
MANN+HUMMEL GmbH, Ludwigsburg;
Dipl. Ing. Christof Mangold,
MANN+HUMMEL GmbH, Ludwigsburg;
Dipl.-Ing. Hedwig Schick,
MANN+HUMMEL GmbH, Ludwigsburg.

Turbochargers and superchargers

Dipl.-Ing. Philipp Schäfer.

Exhaust-gas system

Dr. rer. nat. Rolf Jebasinski,
J. Eberspächer GmbH & Co. KG,
Esslingen.

Management for gasoline engines

Management for gasoline engines

Dipl.-Ing. Armin Hassdenteufel;
Dr.-Ing. Henning Heikes.

Cylinder charge

Dr.-Ing. Martin Brandt;
Dr.-Ing. Henning Heikes.
Dipl.-Ing. (FH) Thomas Kortwittenborg;
Dipl.-Ing. Frank Walter.

Fuel supply system

Dipl.-Ing. Timm Hollmann;
Dipl.-Ing. Thomas Herges;
Dipl.-Ing. Johannes Högl;
Dipl.-Ing. Karsten Scholz;
Dipl.-Ing. Michael Kuehn;
Dipl.-Ing. Zdenek Liner;
Dr.-Ing. Thomas Kaiser;
Dipl.-Ing. Serdar Derikesen;
Dipl.-Ing. Uwe Müller;
Dipl.-Ing. (FH) Horst Kirschner.

Mixture formation

Dipl.-Ing. Andreas Binder;
Dipl.-Ing. Anja Melsheimer;
Dipl.-Ing. Markus Gesk;
Dipl.-Ing. Andreas Glaser;
Dr.-Ing. Tilo Landenfeld.

Ignition

Dr.-Ing. Martin Brandt;
Dipl.-Ing. Walter Gollin;
Dipl.-Ing. Werner Häming;
Dipl.-Ing. Tim Skowronek;
Dr.-Ing. Grit Vogt;
Dr.-Ing. Matthias Budde.

Catalytic exhaust-gas treatment

Dipl.-Ing. Klaus Winkler;
Dipl.-Ing. Detlef Heinrich;
Dr.-Ing. Frank Meier.

Motor sport

Dipl.-Red. Ulrich Michelt,
Bosch Engineering GmbH, Abstatt.

Alternative gasoline-engine operation

LPG operation

Dipl.-Ing. Iraklis Avramopoulos,
IAV GmbH, Berlin.

Engines fueled by natural gas

Dipl.-Ing. (FH) Thorsten Allgeier.

Engines fueled by alcohol

Dipl.-Ing. Andreas Posselt.

Management for diesel engines

Management for diesel engines

Dipl.-Ing. Felix Landhäußer.

Low-pressure fuel supply

Dipl.-Ing. (FH) Stefan Kieferle;
Dr.-Ing. Thomas Kaiser;
Dipl.-Ing. Serdar Derikesen.

Common-rail injection system

Dipl.-Ing. Felix Landhäußer;
Dipl.-Ing. (FH) Andreas Rettich;
Dipl.-Ing. Thilo Klam;
Dr.-Ing. Holger Rapp;
Anees Haider Bukhari, B. Eng, MBA.;
Dipl.-Ing. (FH) Herbert Strahberger.

Time-controlled single-cylinder pump system

Dipl.-Ing. (BA) Jürgen Crepin,
ETAS GmbH, Stuttgart.

Diesel-distributor injection pumps

Dipl.-Ing. (BA) Jürgen Crepin,
ETAS GmbH, Stuttgart.

Start-assist systems

Dipl.-Ing. (FH) Steffen Peischl;
Dipl.-Ing. Friedrich Schmid,
Daimler AG, Stuttgart.

Exhaust-gas treatment

Dr. rer. nat. Christoph Osemann;
Dr.-Ing. Hartmut Lüders.

Drivetrain

Conventional drivetrain, power take-up elements

Dipl.-Ing. Jürgen Wafzig, M. Sc.,
ZF Friedrichshafen AG, Friedrichshafen.

Manual transmission

Dipl.-Ing. Jürgen Wafzig, M. Sc.,
ZF Friedrichshafen AG, Friedrichshafen;
Dipl.-Ing. (FH) Thomas Müller.

Differentials, drivetrain with hybrid drives, drivetrain with electric drives

Dipl.-Ing. Jürgen Wafzig, M. Sc.,
ZF Friedrichshafen AG, Friedrichshafen.

Alternative drives

Hybrid drives

Dipl.-Ing. Thomas Huber.

Boost-recuperation system 48 V

Dr.-Ing. Marc Uhl,
Robert Bosch Starter Motors
Generators GmbH.

Electric machines as drives

Dr.-Ing. Marcus Alexander;
Dr.-Ing. Arndt Kelleter;
Dr.-Ing. Stephan Usbeck
Dipl.-Ing. Uwe Knappenberger.

Fuel cells

Dr.-Ing. Gunter Wiedemann;
Dr. rer. nat. Ulrich Gottwick;
Dipl.-Ing. (FH) Jan-Michael Grähn;
Dipl.-Ing. Dipl.-Wirt.-Ing. Nils Kaiser.

Emission-control and diagnosis legislation

Emission-control Legislation

Dr.-Ing. Matthias Tappe;
Dipl.-Ing. Michael Bender.

Exhaust-gas measuring techniques

Dipl.-Phys. Martin-Andreas Drühe;
Felix Reinke, B.Sc;
Dipl.-Ing. Andreas Weiß,
Dipl.-Ing. (BA) Marc Rottner;
Dipl.-Ing. Andreas Kreh;
Dipl.-Ing. Bernd Hinner;
Dr. rer. nat. Matthias Tappe.

Diagnostics

Dr.-Ing. Markus Willimowski;
Dr. Richard Holberg,
Dr. rer. nat. Hauke Wendt;
Dipl.-Ing. Sani Dzeko.

Vehicle physics

Basic terms of automotive engineering

Prof. Dr. rer. nat. Ludger Dragon,
Daimler AG, Sindelfingen.

Motor-vehicle dynamics

Dipl.-Ing. Marc Birk,
Daimler AG, Stuttgart;
Prof. Dr. rer. nat. Ludger Dragon,
Daimler AG, Sindelfingen;
Dr.-Ing. Rupert Niethammer,
Daimler AG, Stuttgart;
Dipl.-Ing. Imre Boros,
Daimler AG, Stuttgart;
Dipl.-Ing. Klaus Wüst,
Daimler AG, Stuttgart.

Aerodynamics

Dipl.-Ing. Michael Preiß,
Dr. Ing. h.c. F. Porsche AG, Weissach.

Vehicle acoustics

Dipl.-Ing. Hans-Martin Gerhard,
Dr. Ing. h.c. F. Porsche AG, Weissach.

Chassis systems

Chassis systems, suspension

Dipl.-Ing. Maciej Foltanski,
Rheinisch-Westfälische
Technische Hochschule (RWTH)
Aachen.

*Shock and vibration absorbers,
wheel suspensions*

Dipl.-Ing. Jörn Lützow,
fka Forschungsgesellschaft
Kraftfahrwesen mbH Aachen.

Wheels

Dipl.-Ing. Martin Lauer,
Daimler AG, Sindelfingen;
Dipl.-Ing. Werner Hann,
Daimler AG, Sindelfingen;
Dipl.-Hdl. Martin Bauknecht,
MAN Nutzfahrzeuge Group.

Tires

Dipl.-Ing. Dirk Vincken,
Agentur für Text&Bild, Eurasburg;
Dipl.-Ing. Reimund Müller,
Michelin Reifenwerke AG & Co.KGaA
Karlsruhe;
Thilo Balcke,
Michelin Reifenwerke AG & Co.KGaA
Karlsruhe;
Dipl.-Ing. Norbert Polzin;
Frank Kleemann,
Trelleborg Sealing Solutions.

Steering

Dipl.-Ing. Peter Brenner,
Robert Bosch Automotive Steering
GmbH Schwäbisch Gmünd.

Brake systems

Dr. rer. nat. Jürgen Bräuninger;
Werner Schneider.

Brake systems for passenger cars

Werner Schneider;
Dipl.-Ing. Isabell Maier;
Dipl.-Ing. Andreas Burg;
Dipl.-Ing. Andreas Mayer.
Dipl.-Ing. Hubertus Wienken,
Dipl.-Ing. Frank Bährle-Miller;
Dipl.-Ing. Bernhard Kant.

Brake systems for commercial vehicles

Werner Schneider;
Dr.-Ing. Falk Hecker,
Knorr-Bremse SfN, Schwieberdingen;
Dipl.-Ing. Frank Schwab,
Knorr-Bremse SfN, Schwieberdingen;
Dr.-Ing. Dirk Huhn,
ZF Friedrichshafen AG, Friedrichshafen.

Wheel brakes

Werner Schneider;
Dipl.-Ing. Andreas Burg;
Dipl.-Ing. Andreas Mayer;
Dr. Christof Gente;
Dr.-Ing. Falk Hecker,
Knorr-Bremse SfN, Schwieberdingen.

Chassis control and active safety

Wheel-slip control systems

Dr.-Ing. Falk Hecker,
Knorr-Bremse SfN, Schwieberdingen;
Dipl.-Ing. Frank Schwab,
Knorr-Bremse SfN, Schwieberdingen.

Driving-dynamics control system

Dr.-Ing. Gero Nenninger;
Dipl.-Ing. (FH) Jochen Wagner;
Dr.-Ing. Falk Hecker,
Knorr-Bremse SfN, Schwieberdingen.

*Integrated driving-dynamics
control systems*

Dr.-Ing. Michael Knoop.

Passive safety

Occupant-protection systems

Dipl.-Ing. Robert Krott.

Integral-safety systems

Dr.-Ing. Dagobert Masur.

Vehicle bodies

vehicle bodies, passenger cars

Dipl.-Ing. Dieter Scheunert,
Daimler AG, Sindelfingen.

Vehicle bodies, commercial vehicles

Dipl.-Ing. Georg Stefan Hagemann,
Daimler AG, Stuttgart.

Lighting equipment

Dipl.-Ing. Doris Boebel,
Automotive Lighting Reutlingen GmbH;
Dipl.-Ing. Gert Langhammer,
Automotive Lighting Reutlingen GmbH;
Dr.-Ing. Michael Hamm,
Audi AG, Ingoldstadt.

Automotive glazing

Dipl.-Kauffr. (FH) Britta Müller,
Saint-Gobain Sekurit International,
Herzogenrath.

Windshield, rear-window and headlamp cleaning systems

Dr.-Ing. Mario Hüsge;
Dipl.-Ing. Florian Hauser.

Vehicle security systems

Locking systems

Dr.-Ing. Ulrich Nass,
Brose Schließsysteme GmbH & Co. KG,
Wuppertal.

Electronic vehicle immobilizer

Dipl.-Ing. (FH) Manuel Wurm,
ZF Friedrichshafen AG, Friedrichshafen.

Theft-deterrent systems

Dipl.-Ing. (FH) U. Götz.

Automotive electrics

Conventional vehicle electrical systems

Dipl.-Ing. Clemens Schmucker;
Dipl.-Ing. Eberhard Schoch;
Dipl.-Ing. Markus Beck.

Vehicle electrical systems for hybrid drives

Dr.-Ing. Jochen Faßnacht.

Starter batteries

Dr. rer. nat. Eberhard Meißner,
Johnson Controls Autobatterie
GmbH & Co.KGaa, Hannover.

Drive batteries

Dr.-Ing. Stefan Benz,
Robert Bosch Battery Systems GmbH,
Stuttgart;
Dr.-Ing. Christian Pankiewicz,
Robert Bosch Battery Systems GmbH,
Stuttgart;
Dr.-Ing. Holger Fink,
Robert Bosch Battery Systems GmbH,
Stuttgart.

Electric machines

Prof. Dr.-Ing. Jürgen Ulm,
Reinhold-Würth-Hochschule,
Künzelsau;
Reinhardt Erli, M. Sc.,
Reinhold-Würth-Hochschule,
Künzelsau.

Alternators

Dipl.-Ing. Andreas Schröder.

Starting systems

Dipl.-Ing. Andreas Schröder.

Wiring harnesses and plug-in connections

Dr.-Ing. Eckhardt Philipp.

Electromagnetic compatibility (EMC)

Dr.-Ing. Wolfgang Pfaff.

Symbols and circuit diagrams

Editorial staff.

Automotive electronics

ECUs

Dipl.-Ing. Martin Kaiser;
Dipl.-Ing. Axel Aue.

Automotive software engineering

Dipl.-Ing. (BA) Jürgen Crepin,
ETAS GmbH, Stuttgart;
Dr. rer. nat. Kai Pinnow,
ETAS GmbH, Stuttgart;
Dipl.-Ing. Jörg Schäuffele,
ETAS GmbH, Stuttgart.

Automotive networking, buses in motor vehicles

Dr. rer. nat. Harald Weiler;
Dr.-Ing. Tobias Lorenz;
Dipl.-Ing. Oliver Prella,
Dipl.-Ing. Dieter Thoss.

Architecture of electronic systems

Dr.-Ing. Ralf Machauer,
Bosch Engineering GmbH, Abstatt;
Andreas Hörtling, M. Sc.,
Bosch Engineering GmbH, Abstatt;
Andreas Ehrhart, M. Sc.,
Bosch Engineering GmbH, Abstatt.

Sensors

Dr.-Ing. Erich Zabler;
Dipl.-Ökon. Frauke Ludmann;
Dr. rer. nat. Peter Spoden;
Dipl.-Ing. (FH) Cyrille Caillié;
Dr.-Ing. Uwe Konzelmann;
Dr.-Ing. Tilmann Schmidt-Sandte;
Dr.-Ing. Reinhard Neul;
Dr. Berndt Cramer;
Dipl.-Ing. Dipl.-Wirt.-Ing. Nils Kaiser;
Prof. Dr.-Ing. Peter Knoll;
Dipl.-Ing. Joachim Selinger,
M. Sc. (University of Colorado);
Dr.-Ing. Jan Sparbert.

Mechatronics

Dipl.-Ing. Hans-Martin Heinkel;
Dr.-Ing. Klaus-Georg Bürger.

Comfort and convenience

Passenger-compartment climate control

Dipl.-Ing. Peter Kroner,
MAHLE Behr GmbH & Co. KG,
Stuttgart;
Dipl.-Ing. (FH) Thomas Feith,
MAHLE Behr GmbH & Co. KG,
Stuttgart;
Dipl.-Ing. Günter Eberspach,
Eberspächer Climate Control Systems
GmbH & Co. KG, Esslingen;
Ulrich Karl Weber,
Eberspächer Climate Control Systems
GmbH & Co. KG, Esslingen;
Dipl.-Betriebsw. Marcel Bonnet.

*Comfort and convenience systems
in the door and roof areas*

Dipl.-Ing. (FH) Walter Haußecker;
Dipl.-Ing. (FH) Siegfried Reichmann.

*Comfort and convenience functions
in the passenger compartment*

Mattias Hallor, M. Sc.

**User interfaces,
telematics and multimedia**

Display and control

Prof. Dr.-Ing. Peter Knoll.

*Radio and TV reception in motor
vehicles*

Dr.-Ing. Jens Passoke.

Traffic telematics

Dr.-Ing. Michael Weilkes.

Driver-assistance systems

Driver-assistance systems

Dr.-Ing. Thomas Michalke;
Dr.-Ing. Frank Niewels;
Dipl.-Math. (FH) Thomas Lich;
Dr.-Ing. Thomas Maurer.

Computer vision

Dr.-Ing. Wolfgang Niehsen.

Vehicle navigation

Dipl.-Ing. Ernst Peter Neukirchner.

Night-vision systems

Prof. Dr.-Ing. Peter Knoll.

Parking and maneuvering systems

Prof. Dr.-Ing. Peter Knoll.

Adaptive Cruise Control

Dipl.-Ing. Gernot Schröder;
Prof. Dr.-Ing. Peter Knoll.

*Lane assistance and lane change
assistant*

Dr.-Ing. Thomas Michalke;
Dr. rer. nat. Lutz Bürkle.

Emergency braking systems

Dr.-Ing. Thomas Gussner;
Dr.-Ing. Steffen Knoop;
Dr.-Ing. Falk Hecker,
Knorr-Bremse SñN, Schwieberdingen.

Intersection assistant

Dr. rer. nat. Wolfgang Branz;
Dr.-Ing. Rüdiger Jordan.

Active pedestrian protection

Dr.-Ing. Thomas Gussner;
Dr.-Ing. Steffen Knoop.

High-beam assistant

Dipl.-Ing. Doris Boebel,
Automotive Lighting Reutlingen GmbH;
Dipl.-Ing. (FH) Bernd Dreier,
Automotive Lighting Reutlingen GmbH.

Future of automated driving

Dr.-Ing. Frank Niewels;
Dr.-Ing. Rüdiger Jordan.

Quantities and units

In order to be able to express the values of physical quantities, a system of units is required which serves as the yardstick for every measurement. The value of the physical quantity is taken as the product of numerical value and unit. Such a system of units is the SI system, which was established in 1960 by the 11th General Conference on Weights and Measures (Conférence Générale des Poids et Mesures, CGPM). It has since then been adopted by over 50 countries.

In Germany the management of units is by law under the control of the Physikalisch-Technische Bundesanstalt (PTB, based in Braunschweig), Germany's national metrology institute. International responsibility lies with the International Bureau of Weights and Measures (BIPM, Bureau International des Poids et Mesures) in Sèvres near Paris.

Further units which are to this day commonly used (e.g. liter, metric ton, hour, degree Celsius) are in Germany likewise permitted by law and are mentioned here as such.

Units permitted in other countries (e.g. inch, ounce, degree Fahrenheit) or obsolete units are discussed in a separate section.

SI units

SI means "Système International d'Unités" (International System of Units). The system is laid down in ISO 80000 [1] (ISO: International Organization for Standardization) and for Germany in DIN 1301 [2] (DIN: German Institute for Standardization).

Table 1 lists the seven base SI units.

Definitions of the base SI units

The primary definitions of the base SI units are formulated in French. The translation of the original French text is shown in the following in italics (see [3]).

1. Length

The meter is defined as the distance which light travels in a vacuum in 1/299,792,458 seconds (17th CGPM, 1983).

The meter is therefore defined using the speed of light in a vacuum ($c = 299,792,458$ m/s) and no longer by the wavelength of the radiation emitted by the krypton nuclide ^{86}Kr . The meter was originally defined as the forty-millionth part of a terrestrial meridian (standard meter in Paris, 1875).

2. Mass

The kilogram is the unit of mass; it is equal to the mass of the international kilogram prototype (1st CGPM, 1889 and 3rd CGPM, 1901).

This prototype of a platinum-iridium alloy is kept in the BIPM. The national prototype for Germany is kept at the Physikalisch-Technische Bundesanstalt (PTB) in Braunschweig.

3. Time

The second is defined as the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of atoms of the nuclide ^{133}Cs (13th CGPM, 1967).

This definition relates to a cesium atom at rest at a temperature of 0 K. It can be reproduced much more accurately than the earlier astronomical definitions, which

Table 1: Base SI units

Base quantity and symbol	Base SI unit	
	Name	Symbol
Length	<i>l</i> meter	m
Mass	<i>m</i> kilogram	kg
Time	<i>t</i> second	s
Electric current	<i>I</i> ampere	A
Thermodynamic temperature	<i>T</i> kelvin	K
Amount of substance	<i>n</i> mole	mol
Luminous intensity	<i>I</i> candela	cd

were based on the length of a day and a year.

4. Electric current

The ampere is defined as the intensity of a constant electric current which, if maintained in two straight, parallel conductors of infinite length, of negligible circular cross-sections, and placed 1 meter apart in a vacuum, will produce between these conductors a force equal to $2 \cdot 10^{-7}$ newtons (9th CGPM, 1948).

Thus the magnetic field constant is established as $\mu_0 = 4\pi \cdot 10^{-7}$ H/m.

5. Temperature

The kelvin, the unit of thermodynamic temperature, is defined as the fraction 273.16 of the thermodynamic temperature of the triple point of water (10th CGPM, 1954 and 13th CGPM, 1967).

The zero point of the kelvin scale is the absolute temperature zero point. The triple point of pure water, at which all three states of aggregation (solid, liquid, gaseous) exist in equilibrium to each other, is at 273.16 K and a pressure of 611.657 Pa.

6. Amount of substance

6.1. The mole is defined as the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of the carbon nuclide ^{12}C ; its symbol is "mol".

6.2. When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, other particles, or groups of such particles of precisely specified composition (14th CGPM, 1971).

In the definition of the mole it is assumed that the atoms in question are free carbon-12 atoms which are at rest and in the ground state.

7. Luminous intensity

The candela is defined as the luminous intensity in a given direction of a source which emits monochromatic radiation of frequency $540 \cdot 10^{12}$ hertz and of which the radiant intensity in that direction is 1/683 watt per steradian (16th CGPM, 1979).

The "candela" (Latin: candle, emphasis on the 2nd syllable) superseded the "new candle", which is defined by way of the black body radiation for the melting temperature of platinum.

Decimal fractions and multiples of SI units

Decimal fractions and multiples of SI units (base SI units and derived SI units) are denoted by prefixes before the name of the unit (e.g. milligram) or by prefix symbols before the unit symbol (e.g. mg) (Table 2). The prefix symbol is placed without a gap in front of the unit symbol to form a coherent unit.

Prefixes are not used before further units of angle (degree, min, second), of time (minute, hour), and of temperature (degree Celsius).

Table 2: Prefixes for unit of measurements in acc. with DIN 1301 [2]

Prefix	Prefix symbol	Factor	Name of factor
yokto	y	10^{-24}	septillionth
zepto	z	10^{-21}	sextillionth
atto	a	10^{-18}	trillionth
femto	f	10^{-15}	thousand billionth
pico	p	10^{-12}	billionth
nano	n	10^{-9}	thousand millionth
micro	μ	10^{-6}	millionth
milli	m	10^{-3}	thousandth
centi	c	10^{-2}	hundredth
deci	d	10^{-1}	tenth
deca	da	10^1	ten
hecto	h	10^2	hundred
kilo	k	10^3	thousand
mega	M	10^6	million
giga	G	10^9	thousand million ¹
tera	T	10^{12}	billion ¹
peta	P	10^{15}	thousand billion
exa	E	10^{18}	trillion
zetta	Z	10^{21}	sextillion
yotta	Y	10^{24}	septillion

¹ In the USA: $10^9 = 1$ billion, $10^{12} = 1$ trillion.

Derived SI units

SI units are the seven base SI units and all the units derived from them, i.e. which can be represented as a product of powers of the base units. Thus, for example, the unit of force is obtained from Newton's Law $F = ma$ as

$$1 \text{ kg } \frac{\text{m}}{\text{s}^2} = 1 \text{ N (newton).}$$

If the power product contains only the factor 1, the units are referred to as coherently derived units. There are a total of 22 coherently derived units, which like the newton have been assigned their own names (Table 3).

Legal units

The Law on Units in Metrology of 22 February 1985, amended on 29 October 2001, and the related implementing order of 13 December 1985, last amended on 12 July 2008 (as at 2013), stipulate the use of legal units in business and official transactions. Legal units are:

- the SI units,
- decimal fractions and multiples of SI units,
- other permitted units; see the overview on the following pages and [4].

The following tables provide an overview in accordance with DIN 1301 [2].

Table 3: Derived SI units with special names

Quantity	Unit	Unit symbol	Expressed in other SI units
Plane angle	radian	rad	1
Solid angle	steradian	sr	1
Frequency	hertz	Hz	1 Hz = 1/s
Force	newton	N	1 N = 1 kg m/s ² = 1 J/m
Pressure	pascal	Pa	1 Pa = 1 N/m ²
Energy, work, quantity of heat	joule	J	1 J = 1 Nm = 1 Ws
Power	watt	W	1 W = 1 J/s = 1 VA
Celsius temperature	degree Celsius	°C	
Voltage	volt	V	1 V = 1 W/A
Electrical conductance	siemens	S	1 S = 1 A/V = 1/Ω
Electrical resistivity	ohm	Ω	1 Ω = 1 V/A
Electric charge	coulomb	C	1 C = 1 A s
Electrical capacitance	farad	F	1 F = 1 C/V
Inductance	henry	H	1 H = 1 Wb/A
Magnetic flux	weber	Wb	1 Wb = 1 Vs
Magnetic flux density, induction	tesla	T	1 T = 1 Wb/m ²
Luminous flux	lumen	lm	1 lm = 1 cd sr
Luminous intensity	lux	lx	1 lx = 1 lm/m ²
Radioactivity	becquerel	Bq	1 Bq = 1/s
Absorbed dose	gray	Gy	1 Gy = 1 J/kg
Dose equivalent	sievert	Sv	1 Sv = 1 J/kg
Catalytic activity	katal	kat	1 kat = 1 mol/s

Table 4: Legal units

Quantity and symbol	Legal units			Relationship	Remarks and units not to be used, incl. their conversion
	SI	Others	Name		

1. Length, area, volume

Length	l	m		meter	
Area	A	m^2		square meter	
			a	are	$1 \text{ a} = 100 \text{ m}^2$
			ha	hectare	$1 \text{ ha} = 100 \text{ a} = 10^4 \text{ m}^2$
Volume	V	m^3		cubic meter	
			l, L	liter	$1 \text{ l} = 1 \text{ L} = 1 \text{ dm}^3$

2. Angle

(Plane) angle ¹	α, β etc.	rad		radian	$1 \text{ rad} = \frac{1 \text{ m ar}}{1 \text{ m radius}}$	1^g (centesimal degree) = 1 gon
			°	degree	$1^\circ = \frac{\pi}{180} \text{ rad}$	1^c (centesimal minute) = 10^{-2} gon
			'	minute	$1^\circ = 60' = 3,600''$	1^{cc} (centesimal second) = 10^{-4} gon
			"	second		
			gon	gon	$1 \text{ gon} = \frac{\pi}{200} \text{ rad}$	
Solid angle ¹	Ω	sr		steradian	$1 \text{ sr} = \frac{1 \text{ m}^2 \text{ spherical surface}}{1 \text{ m}^2 \text{ sphere radius}^2}$	

¹ The units rad and sr can be replaced by the numeral 1 in calculations.

28 Basic principles

Quantity and symbol	Legal units			Relationship	Remarks and units not to be used, incl. their conversion
	SI	Others	Name		

3. Mass

Mass (weight) ¹	m	kg		kilogram	
			g	gram	$1\text{ g} = 10^{-3}\text{ kg}$
			t	metric ton	$1\text{ t} = 10^3\text{ kg}$
Density	ρ	kg/m ³			Weight of unit volume γ (kp/dm ³ or p/cm ³) $\gamma = \rho \cdot g$
Moment of inertia (mass moment, 2nd order)	J	kg · m ²		$J = m \cdot r^2$ r = radius of gyration	Flywheel effect $G \cdot D^2$ in kp · m ² $D = 2r$, $G = m \cdot g$ $G \cdot D^2 = 4J \cdot g$

4. Time quantities

Time Duration Interval ²	t	s		second	
			min	minute	$1\text{ min} = 60\text{ s}$
			h	hour	$1\text{ h} = 60\text{ min}$
			d	day	$1\text{ d} = 24\text{ h}$
			a	year	
Frequency	f	Hz		hertz	$1\text{ Hz} = 1/\text{s}$
Rotational speed (rotational frequency)	n	s ⁻¹			$1\text{ s}^{-1} = 1/\text{s}$
			min ⁻¹ 1/min		$1\text{ min}^{-1} = 1/\text{min} = (1/60)\text{ s}^{-1}$
Angular frequency	ω	s ⁻¹			$\omega = 2\pi f$
Velocity	v	m/s	km/h		$1\text{ km/h} = (1/3.6)\text{ m/s}$
Acceleration ⁴	a	m/s ²			Normal acceleration of free fall $g \approx 9.80665\text{ m/s}^2$
Angular velocity ³	ω	rad/s			
Angular acceleration ³	α	rad/s ²			

¹ The term "weight" is ambiguous in everyday usage; it is used to denote mass as well as weight (DIN 1305 [5]).

² Clock times: h, min, s written as superscripts. Example: 3^h 25^m 6^s.

³ The unit rad can be replaced by the numeral 1 in calculations.

⁴ Acceleration is sometimes expressed in m/s² as a multiple of gravitational acceleration g .

Quantity and symbol	Legal units			Relationship	Remarks and units not to be used, incl. their conversion
	SI	Others	Name		

5. Force, energy, power

Force	F	N		newton	$1 \text{ N} = 1 \text{ kg} \cdot \text{m/s}^2$	1 kp (kilopond) $\approx 9.80665 \text{ N}$
Force due to weight	G	N				
Impulse	p	Ns			$1 \text{ Ns} = 1 \text{ kg} \cdot \text{m/s}$	
Pressure, gen.	p	Pa		pascal	$1 \text{ Pa} = 1 \text{ N/m}^2$	1 at (techn. atmosphere) $= 1 \text{ kp/cm}^2$ $\approx 0.980665 \text{ bar}$ $p \approx 1.01325 \text{ bar}$ $\approx 1,013.25 \text{ hPa}$ (standard value of air pressure)
			bar	bar	$1 \text{ bar} = 10^5 \text{ Pa}$	
Mechanical stress	σ, τ	N/m^2			$1 \text{ N/m}^2 = 1 \text{ Pa}$	$1 \text{ kp/m}^2 \approx 9.80665 \text{ N/m}^2$
			N/mm^2		$1 \text{ N/mm}^2 = 1 \text{ MPa}$	
Hardness		Brinell and Vickers hardness are no longer given in kp/mm^2 . Instead, an abbreviation of the relevant hardness is written as the unit after the numerical value used previously (including an indication of the test force, etc., where applicable).				Examples: Previously Now $\text{HB} = 350 \text{ kp/mm}^2$ 350 HB HV30 $= 720 \text{ kp/mm}^2$ 720 HV30 $\text{HRC} = 60$ 60 HRC
Energy	E	J		joule	$1 \text{ J} = 1 \text{ Nm} = 1 \text{ Ws}$ $= 1 \text{ kg} \cdot \text{m}^2/\text{s}^2$	1 kcal (kilocalorie) $= 4.1868 \text{ kJ}$
Work	W					
Heat, quantity of heat ¹	Q		Ws	watt-second	$1 \text{ Ws} = 1 \text{ J}$	
			kWh	kilowatt-hour	$1 \text{ kWh} = 3.6 \text{ MJ}$	
			eV	electron-volt	$1 \text{ eV} \approx 1.60219 \cdot 10^{-19} \text{ J}$	
Torque	M	Nm		newton meter		$1 \text{ kp} \cdot \text{m}$ (kilopondmeter) $\approx 9.80665 \text{ Nm}$
Power	P	W		watt	$1 \text{ W} = 1 \text{ J/s} = 1 \text{ Nm/s}$	$1 \text{ kp} \cdot \text{m/s} \approx 9.80665 \text{ W}$ 1 PS (horsepower) $\approx 0.7355 \text{ kW}$
Heat flow	$\dot{Q}$					
Radiated power	Φ					
Apparent power	P_s	VA		volt-ampere	$1 \text{ VA} = 1 \text{ W}$	
Reactive power	P_q	var		var	$1 \text{ var} = 1 \text{ W}$	

¹ The quantity of heat is expressed in joules.

30 Basic principles

Quantity and symbol	Legal units			Relationship	Remarks and units not to be used, incl. their conversion
	SI	Others	Name		

6. Viscosimetric quantities

Dynamic viscosity	η	Pa · s		pascal-second	$1 \text{ Pa} \cdot \text{s} = 1 \text{ N} \cdot \text{s} / \text{m}^2 = 1 \text{ kg} / (\text{s} \cdot \text{m})$	$1 \text{ P (poise)} = 0.1 \text{ Pa} \cdot \text{s}$
Kinematic viscosity	ν	m ² /s			$1 \text{ m}^2 / \text{s} = 1 \text{ Pa} \cdot \text{s} / (\text{kg} / \text{m}^3)$	$1 \text{ St (stokes)} = 10^{-4} \text{ m}^2 / \text{s}$

7. Temperature and heat

Temperature	T	K		kelvin	$\vartheta = (T - 273.15 \text{ K}) \frac{^{\circ}\text{C}}{\text{K}}$	
	ϑ			degree Celsius		
Temperature difference	ΔT	K		kelvin	$1 \text{ K} = 1 ^{\circ}\text{C}$	In the case of composite units, express temperature differences in K
	$\Delta \vartheta$			degree Celsius		

For quantity of heat and heat flow, refer to 5

Specific heat capacity (spec. heat)	c	J/(kg · K)				$1 \text{ kcal} / (\text{kg} \cdot \text{K}) = 4.1868 \text{ kJ} / (\text{kg} \cdot \text{K})$
Molar heat capacity	C	J/(mol · K)				
Thermal conductivity	λ	W/(m · K)				$1 \text{ kcal} / (\text{m} \cdot \text{h} \cdot \text{K}) = 1.163 \text{ W} / (\text{m} \cdot \text{K})$

8. Electrical quantities

Electric current	I	A		ampere		
Electric potential (voltage)	U	V		volt	$1 \text{ V} = 1 \text{ W} / \text{A}$	
Electrical conductance: Conductance Susceptance Admittance	G B Y	S		siemens	$1 \text{ S} = 1 \text{ A} / \text{V} = 1 / \Omega$	
Electrical resistance: Resistance Reactance Impedance	R X Z	Ω		ohm	$1 \Omega = 1 / \text{S} = 1 \text{ V} / \text{A}$	
Quantity of electricity, electric charge	Q	C		coulomb	$1 \text{ C} = 1 \text{ A s}$	
			Ah	ampere-hour	$1 \text{ Ah} = 3,600 \text{ C}$	
Electrical capacitance	C	F		farad	$1 \text{ F} = 1 \text{ C} / \text{V}$	

Quantity and symbol	Legal units			Relationship	Remarks and units not to be used, incl. their conversion
	SI	Others	Name		
Electrical flux density, displacement D	C/m ²			1 C/m ² = 1 As/m ²	
Electric field strength E	V/m			1 V/m = 1 W/(Am)	

9. Magnetic quantities

Inductance L	H		henry	1 H = 1 Wb/A	
Magnetic flux Φ	Wb		weber	1 Wb = 1 Vs	1 M (maxwell) = 10 ⁻⁸ Wb
Magnetic flux density, induction B	T		tesla	1 T = 1 Wb/m ²	1 G (gauss) = 10 ⁻⁴ T
Magnetic field strength H	A/m			1 A/m = 1 N/Wb	1 Oe (oersted) = $\frac{10^3}{(4\pi)} \frac{\text{A}}{\text{m}}$

10. Photometric quantities

Luminous intensity I	Cd		candela		
Luminance L	cd/m ²				1 sb (stilb) = 10 ⁴ cd/m ² 1 asb (apostilb) = 1/π cd/m ²
Luminous flux Φ	lm		lumen	1 lm = 1 cd sr (sr = steradian)	
Illuminance E	lx		lux	1 lx = 1 lm/m ²	

11. Acoustic quantities

Sound pressure p	Pa		pascal		
Sound intensity I	W/m ²				
Sound pressure level L_p	Np		neper	1 Np = 1 (non-dimensional)	$L_p = \ln(p_1/p_2)$ Np $= L_1 = 0.5 \ln(I_1/I_2)$ Np
Sound intensity level L_I		dB	decibel	1 dB = $\frac{1}{20} \ln 10$ Np ≈ 0.1151 Np	$L_p = 20 \lg(p_1/p_2)$ dB $= L_1 = 10 \lg(I_1/I_2)$ dB
		B	bel	1 B = 10 dB ≈ 1.151 Np	At $f = 1,000$ Hz the (physiological) loudness level is measured in "phon": 1 phon = 1 dB and the loudness in "sone": 1 sone = 40 phon

32 Basic principles

Quantity and symbol	Legal units			Relationship	Remarks and units not to be used, incl. their conversion
	SI	Others	Name		
Sound pressure level Sound power level A-weighted	L_{pA} L_{WA}	dB (A)			Frequency-dependent weighting tuned to the human ear at 20 to 40 phon

12. Quantities used in atom physics and other fields

Energy	W		eV	electron-volt	$1\text{ eV} \approx 1.60219 \cdot 10^{-19}\text{ J}$ $1\text{ MeV} = 10^6\text{ eV}$	
Activity of a radioactive substance	A	Bq		becquerel	$1\text{ Bq} = 1\text{ s}^{-1}$	$1\text{ Ci (curie)} = 3.7 \cdot 10^{10}\text{ Bq}$
Absorbed dose	D	Gy		gray	$1\text{ Gy} = 1\text{ J/kg}$	$1\text{ rd (rad)} = 10^{-2}\text{ Gy}$
Dose equivalent	D_q	Sv		sievert	$1\text{ Sv} = 1\text{ J/kg}$	$1\text{ rem} = 10^{-2}\text{ Sv}$
Absorbed dose rate	$\dot{D}$				$1\text{ Gy/s} = 1\text{ W/kg}$	$1\text{ rd/s} = 10^{-2}\text{ Gy/s}$
Ion dose	J	C/kg				$1\text{ R (roentgen)} = 258 \cdot 10^{-6}\text{ C/kg}$
Ion dose rate	$\dot{J}$	A/kg				
Amount of substance	n	mol		mole		
Catalytic activity		kat		katal	$1\text{ kat} = 1\text{ mol/s}$	

Further units

Table 5: Conversion of units

<i>Units of length</i>	
<i>Name</i>	<i>Conversion</i>
micron	$1 \mu = 1 \mu\text{m}$
typographical point	$1 \text{ p} = 0.37607 \text{ mm}$
inch	$1 \text{ in} = 25.4 \text{ mm}$
foot	$1 \text{ ft} = 12 \text{ in} = 0.3048 \text{ m}$
yard	$1 \text{ yd} = 3 \text{ ft} = 0.9144 \text{ m}$
mile	$1 \text{ mile} = 1,760 \text{ yd}$ $= 1.6093 \text{ km}$
nautical mile (international mile)	$1 \text{ NM} = 1 \text{ sm} = 1.852 \text{ km}$ $(\approx 1' \text{ of degree of longitude})$

Anglo-American units of length:

microinch	$1 \mu\text{in} = 0.0254 \mu\text{m}$
milliinch	$1 \text{ mil} = 0.0254 \text{ mm}$
link	$1 \text{ link} = 201.17 \text{ mm}$
rod	$1 \text{ rod} = 1 \text{ pole} = 1 \text{ perch}$ $= 5.5 \text{ yd} = 5.0292 \text{ m}$
fathom	$1 \text{ fathom} = 2 \text{ yd}$ $= 1.8288 \text{ m}$
chain	$1 \text{ chain} = 22 \text{ yd}$ $= 20.1168 \text{ m}$
furlong	$1 \text{ furlong} = 220 \text{ yd}$ $= 201.168 \text{ m}$

Units of area

<i>Name</i>	<i>Conversion</i>
square inch (sq in)	$1 \text{ in}^2 = 6.4516 \text{ cm}^2$
square foot (sq ft)	$1 \text{ ft}^2 = 144 \text{ in}^2$ $= 0.0929 \text{ m}^2$
square yard (sq yd)	$1 \text{ yd}^2 = 9 \text{ ft}^2 = 0.8361 \text{ m}^2$
acre	$1 \text{ ac} = 4,840 \text{ yd}^2$ $= 4,046.9 \text{ m}^2$
square mile (sq mile)	$1 \text{ mile}^2 = 640 \text{ acre}$ $= 2.59 \text{ km}^2$
barn	$1 \text{ b} = 10^{-28} \text{ m}^2$

Units of volume

<i>Name</i>	<i>Conversion</i>
cubic inch (cu in)	$1 \text{ in}^3 = 16.3871 \text{ cm}^3$
cubic foot (cu ft)	$1 \text{ ft}^3 = 1,728 \text{ in}^3$ $= 0.02832 \text{ m}^3$
cubic yard (cu yd)	$1 \text{ yd}^3 = 27 \text{ ft}^3$ $= 0.76456 \text{ m}^3$

Further units in the United Kingdom (UK):

fluid ounce	$1 \text{ fl oz} = 0.028413 \text{ l}$
pint	$1 \text{ pt} = 0.56826 \text{ l}$
quart	$1 \text{ qt} = 2 \text{ pt} = 1.13652 \text{ l}$
gallon	$1 \text{ gal} = 4 \text{ qt} = 4.5461 \text{ l}$
barrel (crude oil)	$1 \text{ bbl} = 35 \text{ gal} = 159.1 \text{ l}$
barrel (other liquids)	$1 \text{ bbl} = 36 \text{ gal} = 163.6 \text{ l}$

Further units in the United States (US):

fluid ounce	$1 \text{ fl oz} = 0.029574 \text{ l}$
liquid pint	$1 \text{ liq pt} = 0.47318 \text{ l}$
liquid quart	$1 \text{ liq qt} = 2 \text{ liq pt}$ $= 0.94635 \text{ l}$
gallon	$1 \text{ gal} = 4 \text{ liq qt} = 3.7854 \text{ l}$
liquid barrel	$1 \text{ liq bbl} = 31.5 \text{ gal}$ $= 119.24 \text{ l}$
barrel petroleum	$1 \text{ barrel petroleum}$ $= 42 \text{ gal} = 158.99 \text{ l}$ (for crude oil)

Velocities

<i>Name</i>	<i>Conversion</i>
kilometer per hour	$1 \text{ km/h} = (1/3.6) \text{ m/s}$ $\approx 0.2778 \text{ m/s}$
miles per hour	$1 \text{ mile/h} = 1.6093 \text{ km/h}$
knot	$1 \text{ kn} = 1 \text{ NM/h}$ $= 1 \text{ sm/h}$ $= 1.852 \text{ km/h}$ $= 0.5144 \text{ m/s}$
Mach number Ma	The Mach number Ma is the quotient of velocity and velocity of sound.

34 Basic principles

Units of mass

Name	Conversion
Pfund (pound)	1 Pfund = 0.5 kg
Zentner (metric hundredweight)	1 Ztr = 50 kg
Doppelzentner (metric quintal)	1 dz = 100 kg
Gamma	1 γ = 1 μ g
metric carat (for precious stones only)	1 Kt = 0.2 g
atomic unit of mass	1 u = $1.6606 \cdot 10^{-27}$ kg
grain	1 gr = 64.79891 mg
pennyweight	1 dwt = 24 gr = 1.5552 g
dram	1 dr = 1.77184 g
ounce	1 oz = 16 dram = 28.3495 g
troy ounce	1 oz tr (US) = 1 oz tr (UK) = 31.1035 g
pound	1 lb = 16 oz = 453.592 g
stone (UK)	1 st = 14 lb = 6.35 kg
quarter (UK)	1 qr = 28 lb = 12.7 kg
slug (mass accelerated by 1 ft/s ² when 1 lbf is exerted on it)	1 slug = 14.4939 kg
hundredweight (US)	1 cwt = 1 cwt sh (short cwt) = 1 quintal = 100 lb = 45.3592 kg
hundredweight (UK)	1 cwt = 1 cwt l (long cwt) = 112 lb = 50.8023 kg
ton (US)	1 ton (US) = 1 tn sh (short ton) = 0.90718 t
ton (UK)	1 ton (UK) = 1 tn l (long ton) = 1 ton dw (ton deadweight) = 1.01605 t
–	1 t dw = 1 t

Density

Name	Conversion
	1 lb/ft ³ = 16.018 kg/m ³
	1 lb/gal (UK) = 99.776 kg/m ³
	1 lb/gal (US) = 119.83 kg/m ³

Areometer degrees n are a measure of the density ρ of a liquid relative to the density of water at 15 °C.

$$n = 144.3(\rho - 1 \text{ kg/l})/\rho^\circ \text{Bé (degree Baumé)}$$

$$n = (141.5 \text{ kg/l} - 131.5 \cdot \rho)/\rho^\circ \text{API}$$

(API: American Petroleum Institute)

Units of force

Name	Conversion
pond	1 p $\approx$ 9.80665 mN
kilopond	1 kp $\approx$ 9.80665 N
dyne	1 dyn = 10^{-5} N
sthène (French)	1 sn = 1 kN
poundforce	1 lbf = 4.44822 N
poundal (force which accelerates a mass of 1 lb by 1 ft/s ²)	1 pdl = 0.138255 N

Units of pressure and stress

Name	Conversion
microbar	1 μ bar = 0.1 Pa
millibar	1 mbar = 1 hPa = 100 Pa
bar	1 bar = 10^5 Pa
–	1 kp/mm ² $\approx 9.80665 \cdot 10^6$ Pa
technical atmosphere	1 at = 1 kp/cm ² $\approx 9.80665 \cdot 10^4$ Pa
physical atmosphere	1 atm = $1.01325 \cdot 10^5$ Pa
torr	1 torr = 1 mmHg (mercury column) = 133.322 Pa
–	1 mm water column = 1 kp/m ² ≈ 9.80665 Pa

The following units are derived from the technical atmosphere:

p_{abs} (absolute pressure): 1 ata

p_{amb} (ambient pressure)

$p_e = p_{\text{abs}} - p_{\text{amb}}$
(pressure above atmospheric):

1 atpaa for $p_{\text{abs}} > p_{\text{amb}}$

$p_e = p_{\text{amb}} - p_{\text{abs}}$
(pressure below atmospheric):

1 atpba for $p_{\text{abs}} < p_{\text{amb}}$

Anglo-American and other units:

poundforce per square inch	1 lbf/in ² = 1 psi = 6894.76 Pa
poundforce per square foot	1 lbf/ft ² = 1 psf = 47.8803 Pa
tonforce per square inch (UK)	1 tonf/in ² = 1.54443 · 10 ⁷ Pa
poundal per square foot	1 pdl/ft ² = 1.48816 Pa
barye (French)	1 barye = 0.1 Pa
pièce (French)	1 pz = 1 sn/m ² (sthène/m ²) = 1,000 Pa

Units of energy

Name	Conversion
erg	1 erg = 10 ⁻⁷ J
calorie ¹	1 cal = 4.1868 J
	1 kp · m = 9.80665 J
	1 HP · h = 2.6478 · 10 ⁶ J
therm	1 therm = 105.50 · 10 ⁶ J

Anglo-American and other units:

inch ounceforce	1 in ozf = 7.062 mJ
foot poundal	1 ft pdl = 0.04214 J
inch poundforce	1 in lbf = 0.11299 J
foot poundforce	1 ft lbf = 1.35582 J
British thermal unit ²	1 Btu = 1,055.06 J
therm	1 therm = 10 ⁵ Btu
horsepower hour	1 hp · h = 2.685 · 10 ⁶ J
thermie (French)	1 thermie = 1,000 frigories = 1,000 kcal = 41.868 MJ
kg coal equivalents	1 kg CE = 29.3076 MJ (for the calorific value H_u = 7,000 kcal/kg coal)
tonne of coal equivalents	1 t CE = 1,000 kg CE

Units of power

Name	Conversion
kilocalorie per hour	1 kcal/h = 1.163 W
calorie per second	1 cal/s = 4.1868 W
—	1 kp · m/s = 9.80665 W
horsepower, cheval vapeur (French)	1 HP = 1 ch = 735.499 W

Anglo-American units:

—	1 ft · lbf/s = 1.35582 W
horsepower	1 hp = 745.70 W
—	1 Btu/s = 1,055.06 W

Fuel consumption

Name	Conversion
—	1 g/(HP · h) = 1.3596 g/(kWh)
—	1 lb/(hp · h) = 608.277 g/(kWh)
—	1 liq pt/(hp · h) = 634.545 cm ³ /(kWh)

$$x \text{ mile/gal (US)} \triangleq \frac{235.21}{x} \text{ l/100 km}$$

$$x \text{ mile/gal (UK)} \triangleq \frac{282.48}{x} \text{ l/100 km}$$

$$y \text{ l/100 km} \triangleq \frac{235.21}{y} \text{ mile/gal (US)}$$

$$y \text{ l/100 km} \triangleq \frac{282.48}{y} \text{ mile/gal (UK)}$$

¹ The quantity of heat which is required to heat 1 g water from 15 °C to 16 °C.

² The quantity of heat which is required to heat 1 lb water from 63 °F to 64 °F.

Units of temperature

Name	Conversion
K (kelvin)	
°C (degree Celsius)	$\vartheta/^{\circ}\text{C} = T/\text{K} - 273.15$ This scale is based on the freezing point of water at 0 °C and the boiling point at 100 °C, at normal pressure in each case
°F (degree Fahrenheit)	$T_{\text{F}}/^{\circ}\text{F} = 1.8 \cdot T/\text{K} - 459.67$ The freezing point of water is 32 °F, human body temperature is 96 °F
°Ra (degree Rankine)	$T_{\text{Ra}}/^{\circ}\text{Ra} = 1.8 \cdot T/\text{K}$ This scale begins with the absolute temperature zero point at 0 K and uses a graduation like the Fahrenheit
°Re (degree Réaumur)	$T_{\text{Re}}/^{\circ}\text{Re} = 0.8 \cdot (T/\text{K} - 273.15)$ The freezing point of water is 0 °Re, the boiling point is 80 °Re
Temperature differences	$1 \text{ K} \triangleq 1^{\circ}\text{C} \triangleq 1.8^{\circ}\text{F}$ $\triangleq 1.8^{\circ}\text{Ra} \triangleq 0.8^{\circ}\text{Re}$

Units of viscosity (kinematic)

Name	Conversion
	$1 \text{ ft}^2/\text{s} = 0.092903 \text{ m}^2/\text{s}$ Measurement is carried out by viscosimeter in accordance with the standard DIN EN ISO 2431 [6]
A-seconds	Runout time from flow cup
Engler number	Relative runout time from Engler device: $1^{\circ}\text{E} \approx 7.6 \text{ mm}^2/\text{s}$
RI seconds	Runout time from Redwood-I viscosimeter (UK): $1 \text{ R}'' \approx 4.06 \text{ mm}^2/\text{s}$
SU seconds	Runout time from Saybolt Universal viscosimeter (US): $1 \text{ S}'' \approx 4.63 \text{ mm}^2/\text{s}$

Natural constants

It is currently assumed that the physical limits apply equally at any place in our world. These describe the natural constants which both contain physical quantities (e.g. speed of light in vacuum, masses of elementary particles)

Table 5: Natural constants

Avogadro constant N_A	$6.022\,141\,5(10) \cdot 10^{23} \text{ mol}^{-1}$
Faraday constant F	$F = e N_A$ $\approx 96485.3365(21) \frac{\text{C}}{\text{mol}}$
Loschmidt number N_L	$N_L = \frac{N_A}{V_{m0}}$ $\approx 2.6867805(24) \cdot 10^{25} \text{ m}^{-3}$ V_{m0} : Molar volume of an ideal gas under normal conditions
Boltzmann constant k	$1.380\,650\,5(24) \cdot 10^{-23} \frac{\text{J}}{\text{K}}$
elementary charge e	$1.602\,176\,53(14) \cdot 10^{-19} \text{ C}$
electric field constant ϵ_0	$\frac{1}{\mu_0 \cdot c^2}$ $\approx 8.854\,187\,817\,62 \cdot 10^{-12} \frac{\text{F}}{\text{m}}$
magnetic field constant μ_0	$4\pi \cdot 10^{-7} \frac{\text{Vs}}{\text{Am}}$ $\approx 12.566\,370\,614 \cdot 10^{-7} \frac{\text{H}}{\text{m}}$
gravitation constant G	$6.674\,2(10) \cdot 10^{-11} \frac{\text{m}^3}{\text{kg} \cdot \text{s}^2}$
velocity of light c (in vacuum)	$2.997\,924\,58 \cdot 10^8 \frac{\text{m}}{\text{s}}$
atomic unit of mass u	$1.660\,538\,86(28) \cdot 10^{-27} \text{ kg}$
Planck action quantum h	$6.626\,069\,3(11) \cdot 10^{-34} \text{ Js}$
rest mass of electron m_e	$9.109\,382\,6(16) \cdot 10^{-31} \text{ kg}$
rest mass of proton m_p	$1.672\,621\,71(29) \cdot 10^{-27} \text{ kg}$
Stefan-Boltzmann constant σ	$5.670\,400\,(40) \cdot 10^{-8} \frac{\text{W}}{\text{m}^2 \cdot \text{K}^4}$
universal gas constant R	$8.314\,472\,(15) \frac{\text{J}}{\text{mol} \cdot \text{K}}$

and represent the connections between them (e.g. field constants of gravitation, electricity and magnetism). Their values depend on the system of units and must be experimentally determined.

Table 5 lists the most important natural constants (source: [4]). The numbers in parentheses express the inaccuracy of the last two digits.

References

- [1] ISO 80000: Quantities and units (2013).
- [2] DIN 1301: Units – Part 1: Unit names, unit symbols (2010).
- [3] 8th Edition of the SI Brochure, printed in the PTB Reports, 117th year, Volume 2, June 2007, Braunschweig and Berlin.
- [4] PTB leaflet “The legal units in Germany” (2004), Braunschweig.
- [5] DIN 1305: Mass, as weighed value, force, weight force, weight, load; concepts. (1988).
- [6] DIN EN ISO 2431: Paints and varnishes – Determination of flow time by use of flow cups (ISO 2431:2011); German version EN ISO 2431:2011.

Mathematical signs and symbols

+	plus
−	minus
· or ×	multiplied by
: or /	divided by
=	equal to
≈	approximately equal to
≠	not equal to
<	less than
>	greater than
≤	less than or equal to
≥	greater than or equal to
~ or ∝	proportional to
$\sum a_i$	sum over a_i
$\prod a_i$	product of a_i
$n!$	n factorial ($1 \cdot 2 \cdot 3 \dots n$)
Δ	difference or Laplace operator
$\sqrt{\quad}$	root
	parallel to
⊥	perpendicular to
→	approaches
∞	infinity
d/dx	differentiation to x
$\partial/\partial x$	partial differentiation to x
$\int f(x)dx$	integral of $f(x)$

Greek alphabet

A	α	Alpha
B	β	Beta
Γ	γ	Gamma
Δ	δ	Delta
E	ε, ϵ	Epsilon
Z	ζ	Zeta
H	η	Eta
Θ	θ, ϑ	Theta
I	ι	Iota
K	κ, κ	Kappa
Λ	λ	Lambda
M	μ	Mu
N	ν	Nu
Ξ	ξ	Xi
O	$\omicron$	Omicron
Π	π, ϖ	Pi
P	ρ, ϱ	Rho
Σ	σ, ς	Sigma
T	τ	Tau
Υ, γ	υ	Upsilon
Φ	ϕ, φ	Phi
Χ	χ	Chi
Ψ	ψ	Psi
Ω	ω	Omega

Basic principles of mechanics

Rectilinear and rotary motion

Forces, moments and energies are required to move bodies with mass. Any number of motions can be composed of rectilinear motions (translation) and rotary motions (rotation). The most important basic equations for these motions are set out in Table 2. The quantities used are listed in Table 1.

Mass and moment of inertia

Mass m is a property of matter and the cause of the inertia which sets a resistance against a change of velocity (acceleration) in the case of a rectilinear motion. The moment of inertia J (also called mass moment of inertia or rotating mass) causes a resistance in the case of a rotary motion (Table 3).

Path, velocity and acceleration

Path s is a limited distance, velocity v is the path covered during a specific time t . For a rotary motion angular velocity ω is obtained from the angle φ covered during a specific time. A uniform motion exists when velocity v or rotational speed n (or angular velocity ω) is constant. In this case, acceleration (a or α) equals zero.

A body is accelerated when there is a change of velocity. A motion is uniformly accelerated if acceleration is constant. In the case of negative acceleration, the motion is decelerated or braked.

Force and moment

A force accelerates or deforms a body. It is referred to in classic physics as the time rate of change of linear momentum p . If a mass m is rotated at centroidal dis-

Table 1: Symbols and units

Quantity	Unit
A Area	m^2
E Modulus of elasticity	N/mm^2
E_k Kinetic energy	$\text{J} = \text{Nm}$
E_p Potential energy	$\text{J} = \text{Nm}$
E_{rot} Rotational energy	$\text{J} = \text{Nm}$
F Force	N
F_G Weight	N
F_m Mean force during impulse period	N
F_R Frictional force	N
F_U Peripheral force	N
F_z Centrifugal force	N
H Rotational impulse	$\text{Nm} \cdot \text{s} = \text{kg} \cdot \text{m}^2/\text{s}$
I Force impulse	$\text{Ns} = \text{kg} \cdot \text{m/s}$
J Moment of inertia	$\text{kg} \cdot \text{m}^2$
L Angular momentum	$\text{Nm} \cdot \text{s}$
M_t Torque	Nm
$M_{t,m}$ Mean torque during impulse period	Nm
P Power	$\text{W} = \text{Nm/s}$
R_e Yield point	N/mm^2
R_m Tensile strength	N/mm^2
V Volume	m^3
W Work, energy	$\text{J} = \text{Nm}$
a Acceleration	m/s^2

Quantity	Unit
a_z Centrifugal acceleration	m/s^2
d Diameter	m
e Base of natural logarithms ($e \approx 2.781$)	—
g Acceleration of free fall ($g \approx 9.81$)	m/s^2
h Height	m
i Radius of gyration	m
l Length	m
m Mass	kg
n Rotational speed	$1/\text{s}$
p Linear momentum	Ns
p Surface pressure	N/mm^2
r Radius	m
s Length of path	m
t Time	s
v Velocity	m/s
α Angular acceleration	rad/s^2
β Wrap angle	$^\circ$
γ Wedge angle	$^\circ$
ϵ Elongation	$\%$
μ Coefficient of friction	—
ν Transverse contraction	—
ρ Density	kg/m^3
φ Angle of rotation	rad
ω Angular velocity	$1/\text{s}$

tance r about a rotational axis at angular velocity ω , this generates a centrifugal force F_z which acts radially outwards from the center. A force which is applied at distance r from a center of rotation generates a moment. An accelerated rotary motion with rotating mass J causes a moment in the form of an acceleration or braking torque.

Work and energy

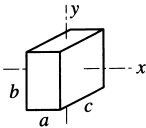
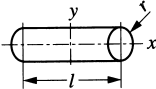
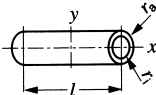
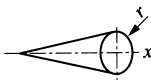
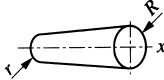
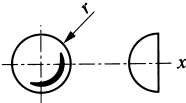
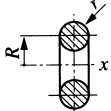
When a force F moves a body by path s , work W is performed which is stored as

energy in this body. Energy is therefore defined in physics as stored work. Conversely, an energy can perform work. A distinction is made in mechanics between kinetic and potential energy. Kinetic energy is the work that must be expended to accelerate a body with mass m or rotating mass J to velocity v or angular velocity ω . Potential energy is the work that must be expended to raise a body to a height h . When a spring is tensioned, potential energy is stored which can perform work again when the spring is released.

Table 2: Rectilinear and rotary motion

Rectilinear motion (translation)	Rotary motion (rotation)
Mass $m = V\rho$	Moment of inertia (Table 3) $J = m i^2$
Path $s = \int v(t) dt$ $s = vt$ [v = const.] $s = \frac{1}{2}at^2$ [a = const.]	Angle $\varphi = \int \omega(t) dt$ $\varphi = \omega t = 2\pi n$ [n = const.] $\varphi = \frac{1}{2}\alpha t^2$ [α = const.]
Velocity $v = ds(t)/dt$ $v = s/t$ [v = const.] $v = at = \sqrt{2as}$ [a = const.]	Angular velocity $\omega = d\varphi(t)/dt$ $\omega = \varphi/t = 2\pi n$ [n = const.] $\omega = \alpha t = \sqrt{2\alpha\varphi}$ [α = const.] Peripheral velocity $v = r\omega$
Acceleration $a = dv(t)/dt$ $a = (v_2 - v_1)/t$ [a = const.]	Angular acceleration $\alpha = d\omega(t)/dt$ $\alpha = (\omega_2 - \omega_1)/t$ [α = const.] Centrifugal acceleration $a_z = r\omega^2$
Force $F = ma$	Torque $M_t = Fr = J\alpha$ Centrifugal force $F_z = mr\omega^2$
Work $W = Fs$	Rotational work $W = M_t\varphi$
Translation energy $E_k = \frac{1}{2}mv^2$	Rotational energy $E_{rot} = \frac{1}{2}J\omega^2$
Potential energy $E_p = F_G h$	
Power $P = dW/dt = Fv$	Power $P = dW/dt = M_t\omega = M_t \cdot 2\pi n$
Linear momentum $p = mv$	Angular momentum $L = J\omega = J \cdot 2\pi n$
Force impulse $I = \Delta p = F_m(t_2 - t_1)$	Rotational impulse $H = M_{t, m}(t_2 - t_1)$

Table 3: Moments of inertia

Type of solid	Moments of inertia (J_x about the x -axis ¹ , J_y about the y -axis ¹)
Right parallelepiped, cuboid 	$J_x = m \frac{b^2 + c^2}{12}$, $J_y = m \frac{a^2 + c^2}{12}$ $J_x = J_y = m \frac{a^2}{6}$ (cube with side length a)
Regular cylinder 	$J_x = m \frac{r^2}{2}$, $J_y = m \frac{r^2 + l^2}{12}$
Hollow regular cylinder 	$J_x = m \frac{r_a^2 + r_i^2}{2}$, $J_y = m \frac{r_a^2 + r_i^2 + \frac{l^2}{3}}{4}$
Circular cone 	$J_x = m \frac{3r^2}{10}$ $J_x = m \frac{r^2}{2}$ Envelope of cone (excluding end base)
Truncated circular cone 	$J_x = m \frac{3(R^5 - r^5)}{10(R^3 - r^3)}$ $J_x = m \frac{(R^2 + r^2)}{2}$ Envelope of cone (excluding end faces)
Sphere and hemisphere 	$J_x = m \frac{2r^2}{5}$ $J_x = m \frac{2r^2}{3}$ Surface area of sphere
Hollow sphere r_a Outer sphere radius r_i Inner sphere radius	$J_x = m \frac{2(r_a^5 - r_i^5)}{5(r_a^3 - r_i^3)}$
Torus 	$J_x = m (R^2 + \frac{3}{4}r^2)$

¹ The moment of inertia for an axis parallel to the x -axis or y -axis at a distance a is:
 $J_A = J_x + m a^2$ or $J_A = J_y + m a^2$.

Conservation of energy

The law of conservation of energy states that the total energy in a closed system is constant. Energy can be neither created nor destroyed, but instead converted between different forms of energy (e.g. kinetic energy to thermal energy) or transmitted from one body to another.

Power

Power is the work performed during a specific period of time. Because every power transfer is associated with loss, the output power is always less than the input power. The ratio of output to input power is called efficiency η and is therefore always less than 1.

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}}.$$

Linear momentum and impulse

Linear momentum p describes the motion of a body with mass, and is calculated as the product of moving mass m and velocity v . Every moving body can transmit its linear momentum during an impulse process to another body, as happens for example in a collision between two vehicles. The force acting on a body gives rise to a change of linear momentum, which is called force impulse I .

During a rotary motion angular momentum L is obtained from the product of rotating mass J and angular velocity ω . A rotational impulse H is generated for example when two disks are jerkily coupled.

Conservation of momentum

The law of conservation of momentum states that the total momentum in a closed system is constant. It follows from this that the total momentum before and after an impulse must be equal.

Statics

Statics is the science of equilibrium on a rigid body. A body is in equilibrium when it is at rest or is moving uniformly or in a straight line. Equilibrium prevails when the sum of the applied forces and moments in all directions is equal to zero.

Plane system of forces

Forces are vectors which are determined by size and direction. They are geometrically (vectorially) added:

$$\vec{F}_{\text{res}} = \vec{F}_1 + \vec{F}_2.$$

Two forces are composed with a parallelogram of forces or a triangle of forces (Figure 1a). If there are several forces, the resultant force F_{res} is determined by means of a polygon of forces (Figure 1b). When the polygon of forces is closed, the system of forces is in equilibrium.

Figure 1: Composition of forces

- a) Parallelogram of forces and triangle of forces,
b) Closed polygon of forces (system in equilibrium).
 F Force,
 F_{res} Resultant force.

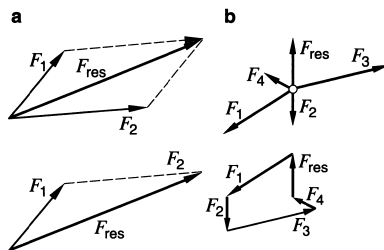
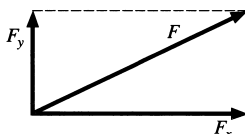


Figure 2. Force resolution

- F Force,
 F_x Force in x -direction,
 F_y Force in y -direction.



A force can also be resolved into components. A resolution into orthogonal components is sensible (Figure 2).

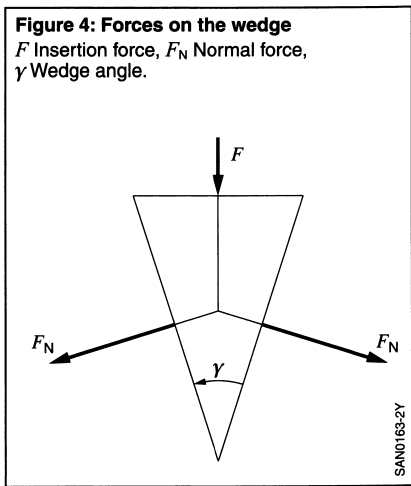
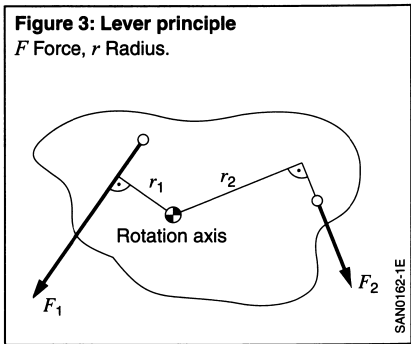
Transmission of force

Mechanical machines for the transmission of force can be reduced to the “lever” and “wedge” principles.

Lever

The lever principle can be derived from the equilibrium condition “sum of the moments is equal to zero”. Disregarding friction, the system in Figure 3 is in equilibrium for the following condition:

$$M_{t1} = M_{t2} \quad \text{or} \quad F_1 r_1 = F_2 r_2 .$$



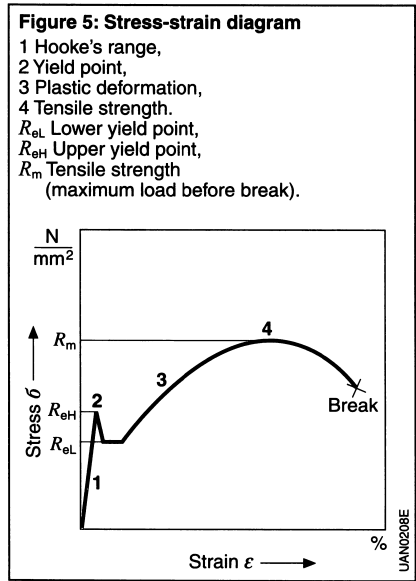
The lever principle is encountered in many applications, ranging from simple pliers, scales and wrenches through gear wheels and belt drives, right down to the connecting rods in piston engines.

Wedge

With the wedge principle, depending on the wedge angle *γ*, small forces (insertion force *F*) can be translated into large normal forces *F_N* (Figure 4). Without allowing for friction, the following applies:

$$F_N = \frac{F}{2 \sin \frac{\gamma}{2}} .$$

Very large application forces can be introduced in the smallest of spaces with the aid of wedges. Examples include wedges in shaft/hub connections and tapered connections for transmitting torques. However, both the screw and the clamping eccentric work according to the classic wedge principle.



Calculation of strength

Hooke's law

In response to an external load a body is deformed and stresses are generated inside the material (Figure 5). Metals have linear-elastic behavior until the yield point is reached. In other words, within this Hooke's range the component assumes its original length again when the load is reduced. The yield point R_e is the limit between elastic and plastic deformation. Above the yield point the material begins to "creep" and then remains permanently deformed. The tensile strength R_m is the maximum load before the component cracks or breaks.

In the case of ductile materials with a pronounced yield point R_e is always specified as the limit for dimensioning. The tensile strength is used for dimensioning only in the case of brittle materials which do not have a pronounced yield point (e.g. gray cast iron).

The linear area is known as Hooke's straight line, in which the ratio of stress σ and strain (elongation) ε is constant. With the proportionality constant E , which is termed the modulus of elasticity, Hooke's law for the mono-axial stress state is:

$$\sigma = E \varepsilon .$$

Strength verification

The aim of strength verification is to dimension components reliably and appropriately for the material involved. Strength verification in accordance with Figure 6 is conducted in four stages:

1. Determination of the external load (forces and moments),
2. Calculation of the existing stress,
3. Choice of the material characteristic value,
4. Comparison of the existing stress with the material characteristic value.

If the external forces and moments are known, the stresses in the component can be calculated in accordance with Table 4. Because tensile, compressive and bending stresses (normal stresses) lie in one plane or act in the same direction, they can be cumulatively superimposed. Even the shear and torsional stresses can be added. However, if normal and shear stresses occur simultaneously in a component, a reduced stress must be created with a strength hypothesis because the material characteristic values were calculated from mono-axial tensile and vibration fatigue tests.

There are different hypotheses to suit the material behavior (ductile or brittle). For ductile materials (most metals) the deformation energy hypothesis is used. For the mono-axial stress state (bar) the reduced stress is calculated in accordance with Table 4.

Surface pressure

Surface pressure is a compressive stress. It is generated when a force F is transmitted from one solid body to another. For effective surface areas pressure p is equal to the ratio of force F to contact surface A .

Pressure on the face of a hole

When journal and bore or shaft and friction bearing are paired, the pressure on the face of a hole is usually used for calculation. Here the load is referred to the projected area A_{proj} (Figure 7):

$$p = \frac{F}{A_{\text{proj}}} = \frac{F}{dl} \leq p_{\text{perm}} .$$

Figure 6: Strength verification

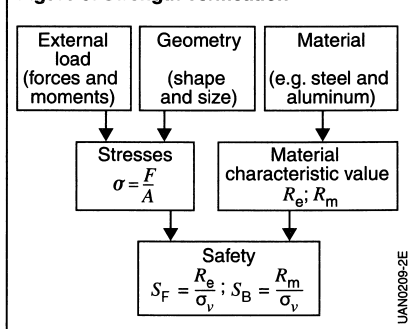
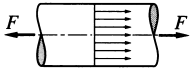
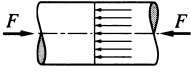
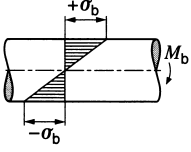
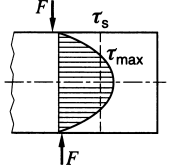
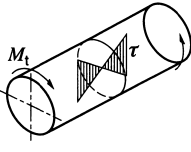
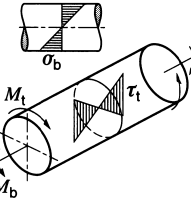


Table 4: Basic equations of strength calculation

Loading	Stress
Tensile stress 	$\sigma_z = \frac{F}{A}$
Compressive stress 	UAN0214Y $\sigma_d = \frac{F}{A}$
Bending stress 	UAN0215Y $\sigma_b = \frac{M_b}{W_b}$ W_b from Table 5
Shear stress 	$\tau_s = \frac{F}{A}$
Torsional stress 	UAN0217Y $\tau_t = \frac{M_t}{W_t}$ W_t from Table 5
Reduced stress 	UAN0218Y $\sigma_v = \sqrt{\sigma^2 + 3\tau^2}$

Hertzian stress

In reality, however, the maximum stress where the surfaces are curved is greater and can be calculated according to Hertz's theory. The maximum Hertzian stress is dependent on the deformation (flattening) of the touching surfaces (Figure 8). This in turn is dependent on the radii, on the modulus of elasticity E and

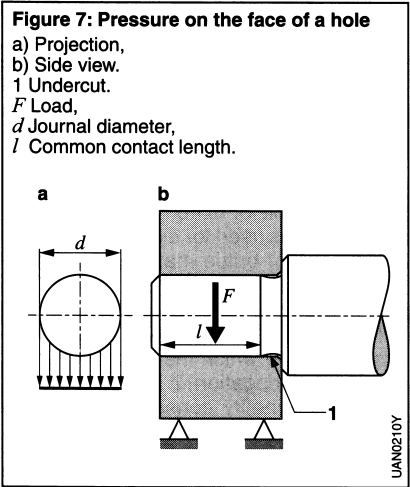
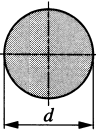
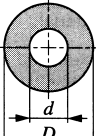
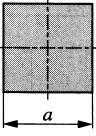


Table 5: Section moduli

Cross-section	Bending	Torsion
	$W_b = \frac{\pi}{32} d^3$	UAN0219Y $W_t = \frac{\pi}{16} d^3$
	$W_b = \frac{\pi}{32} \left(\frac{D^4 - d^4}{D} \right)$	UAN0220Y $W_t = \frac{\pi}{16} \left(\frac{D^4 - d^4}{D} \right)$
	$W_b = \frac{a^3}{6}$	UAN0221Y $W_t = 0.208 a^3$

on the transverse contraction v . The equations for the cases “sphere – sphere” and “cylinder – cylinder” are set out in Table 6. For the special cases “sphere – plane” and “cylinder – plane” radius $r_2 \rightarrow \infty$ or $r = r_1$.

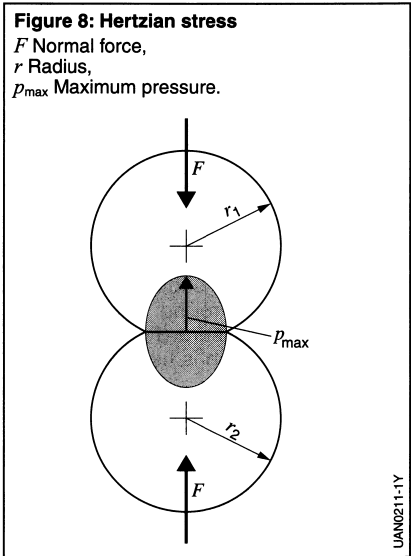


Table 6: Hertzian stress	
Sphere – sphere (localized contact)	Mean <i>E</i> modulus: $E = 2 \cdot \frac{E_1 E_2}{E_1 + E_2}$
$p_{\max} = \frac{1}{\pi} \sqrt{\frac{1.5 F E^2}{r^2 (1 - \nu^2)^2}}$	For the radii: $\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} \rightarrow r = \frac{r_1 r_2}{r_1 + r_2}$
Cylinder – cylinder (line contact)	<i>l</i> Contact length, cylinder <i>v</i> Transverse contraction
$p_{\max} = \sqrt{\frac{F E}{2 \pi r l (1 - \nu^2)}}$	

Friction

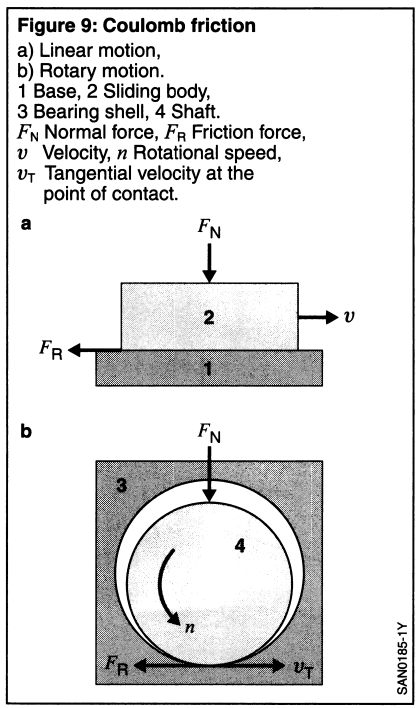
Coulomb friction

When touching bodies move relative to each other, friction acts as mechanical resistance acting in the opposite direction to the motion at velocity v (Figure 9). The force of resistance, known as frictional force F_R , is proportional to the normal force F_N . Static friction exists as long as the external force is less than the frictional force and the body remains at rest. When static friction is overcome, and the body is set in motion, frictional force is governed by Coulomb's law of sliding friction:

$$F_R = \mu F_N.$$

Coefficient of friction

The coefficient of friction μ always denotes a system property and not a material property. Coefficients of friction are, among other things, depending on material pairing (see Table 7), temperature,



surface condition, sliding velocity, the surrounding medium (e.g. water or CO₂, which can be adsorbed by the surface), and the intermediate material (lubricant). For this reason, coefficients of friction always fluctuate between limit values and may have to be calculated experimentally. Static friction is generally greater than sliding friction. In special cases, the friction coefficient can exceed 1, for example with very smooth surfaces where cohesion forces are predominant or with racing tires featuring an adhesion or suction effect.

Friction on the wedge

Allowing for friction, the insertion force is governed by (Figure 10):

$$F = 2 \left(F_N \sin \frac{\gamma}{2} + F_R \cos \frac{\gamma}{2} \right)$$

The normal force is then

$$F_N = \frac{F}{2 \left(\sin \frac{\gamma}{2} + \mu \cos \frac{\gamma}{2} \right)}$$

Rope friction

Motions and moments can be transmitted with elastic, pliable ropes or belts (Figure 11). Sliding friction occurs in the event of relative motion between rope and pulley (e.g. belt brake or bollard with

running rope). Static friction exists when rope and pulley are at rest relative to each other (e.g. belt drive, belt brake as holding brake, bollard with rope at rest). The coefficient of sliding friction μ or the static friction μ_H must be applied accordingly.

According to Euler's rope-friction formula, the force in the pulling strand (taut strand):

$$F_1 = F_2 e^{\mu \beta}$$

The peripheral force F_U or frictional force F_R is calculated as

$$F_U = F_R = F_1 - F_2$$

and the transmittable friction torque as

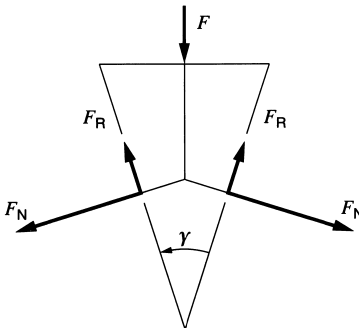
$$M_R = F_R r.$$

Rolling friction

Rolling friction occurs when a ball, a caster or a wheel rolls on a track or roadway. Typical examples of this are the roller bearing, the pairing of wheel flange and rail for railroads or the pairing of tire and roadway for motor vehicles. Both rolling body and base are subjected to elastic deformation during rolling. This generates asymmetrical pressure (Figure 12). The equilibrium condition is used to calculate

Figure 10: Friction on the wedge

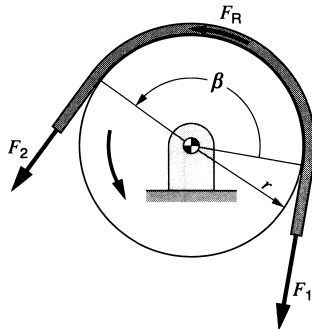
F_N Normal force,
 F_R Frictional force,
 γ Wedge angle.



SAN0164-3Y

Figure 11: Rope friction

F_1 Force in taut strand,
 F_2 Force in slack strand,
 r Radius of pulley,
 β Wrap angle.



UAN0212Y

frictional force F_R for the dry state (i.e. without lubrication):

$$F_R = \frac{x}{R} F_N = \mu_R F_N .$$

Ratio x/R can be calculated as frictional resistance or coefficient of rolling friction μ_R . From this it can be seen that large rolling bodies roll more easily than small rolling bodies. Hard surfaces (roller bearing and railroad) give rise to small deformations and therefore result in very low coefficients of friction; soft surfaces, as in the tire/roadway pairing, on the other hand result in higher coefficients of friction. While frictional resistances of $\mu_R = 0.0015$ can be achieved for ball bearings, the coefficients of friction of a car tire on asphalt are $\mu_R = 0.015$.

References

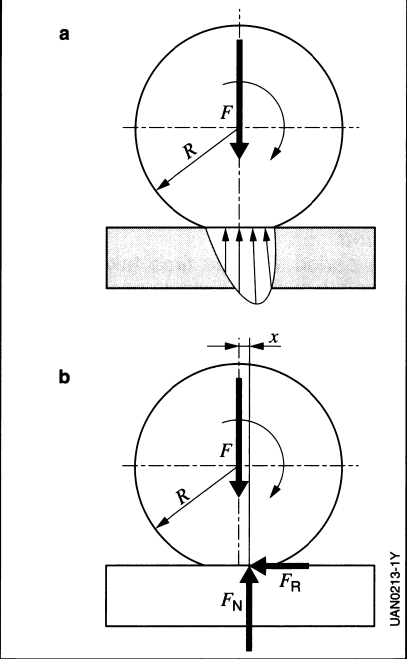
[1] A. Böge: Technische Mechanik. 30th Ed., Verlag Springer Vieweg, 2013.
[2] R.C. Hibbeler: Technische Mechanik 1, 2 und 3, Pearson Studium.
[3] K.-H. Grote, J. Feldhusen: Dubbel – Taschenbuch für den Maschinenbau. 23rd Ed.; Springer-Verlag 2011.

Table 7: Reference values for coefficients of static and sliding friction

Material pair	Coefficient of static friction μ_H		Coefficient of sliding friction μ	
	Dry	Lubricated	Dry	Lubricated
Steel – steel	0.15 – 0.20	0.10	0.10 – 0.15	0.05 – 0.10
Steel – cast iron	0.18 – 0.25	0.10	0.15 – 0.20	0.10
Steel – sintered bronze	0.20 – 0.40	0.08 – 0.13	0.18 – 0.30	0.06 – 0.09
Steel – brake pad	– ¹	– ¹	0.50 – 0.60	0.20 – 0.50
Steel – polyamide	0.60	0.20	0.32 – 0.45	0.10
Steel – ice	0.027	– ¹	0.014	– ¹
Cast iron – cast iron	0.18 – 0.25	0.10	0.15 – 0.20	0.10
Aluminum – aluminum	0.50 – 1.00	– ¹	0.50 – 1.00	– ¹
Wood – wood	0.40 – 0.60	0.20	0.20 – 0.40	0.10
Wood – metal	0.60 – 0.70	0.10	0.40 – 0.50	0.10
Leather – metal	0.50 – 0.60	0.20 – 0.25	0.20 – 0.25	0.12
Car tire on dry asphalt	0.50 – 0.60	– ¹	– ¹	– ¹
Car tire on wet asphalt	0.20 – 0.30	– ¹	– ¹	– ¹
Car tire on icy asphalt	0.10	– ¹	– ¹	– ¹

¹ No practical use.

Figure 12: Rolling friction
a) Asymmetrical pressure distribution,
b) Resultant forces for calculation model.
 F Load, F_N Normal force, F_R Frictional force,
 R Radius of rolling body,
 x Lever arm for moment equilibrium.



UAN0213-1Y

Vibrations and oscillations

Terms

(Symbols and units in Table 1, see also DIN 1311, [1] and [2]).

Vibration or oscillation

A vibration or oscillation is the change in a physical quantity which repeats at more or less regular time intervals and whose direction changes with similar regularity (Figure 1).

Period

The period T is the time taken for one complete cycle of a single oscillation.

Amplitude

The amplitude $\hat{y}$ is the maximum instantaneous value (peak value) of a sinusoidally oscillating physical quantity.

Frequency

The frequency f is the number of oscillations in one second, the reciprocal value of the period of oscillation T .

Angular frequency

The angular frequency ω is 2π times the frequency f .

Particle velocity

Particle velocity v is the instantaneous value of the alternating velocity of a vibrating particle in its direction of vibration. It must not be confused with the velocity of propagation of a traveling wave (e.g. the velocity of sound).

Figure 1: Sinusoidal oscillation
(Quantities, see Table 1)

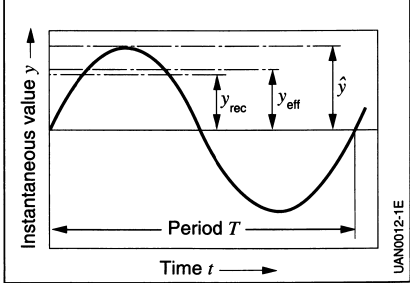


Table 1: Symbols and units

Quantity		Unit
a	Storage coefficient	
b	Damping coefficient	
c	Storage coefficient	
c_s	Spring constant	N/m
c_α	Torsional rigidity	Nm/rad
C	Capacitance	F
f	Frequency	Hz
f_g	Resonant frequency	Hz
Δf	Half-value width	Hz
F	Force	N
F_Q	Excitation function	
I	Current	A
J	Moment of inertia	kg · m ²
L	Inductance	H
m	Mass	kg
M	Torque	Nm
n	Rotational speed	1/min
Q	Charging	C
Q	Resonance sharpness	
r	Damping factor	Ns/m
r_α	Rotational damping coefficient	Ns · m
R	Ohmic resistance	Ω
t	Time	s
T	Period	s
U	Voltage	V
v	Particle velocity	m/s
x	Travel/displacement	m
y	Instantaneous value	
$\hat{y}$	Amplitude	
$\dot{y}$ ($\ddot{y}$)	Single (double) derivative with respect to time	
y_{rec}	Rectification value	
y_{eff}	Effective value	
α	Angle	rad
δ	Decay coefficient	1/s
Λ	Logarithmic decrement	
ω	Angular velocity	rad/s
ω	Angular frequency	1/s
Ω	Excitation-circuit frequency	1/s
D	Damping ratio	
D_{opt}	Optimum damping ratio	
Subscripts:		
0	Undamped	
d	Damped	
T	Absorber	
U	Base support	
G	Machine	

Fourier series

Every periodic function, which is piecewise monotonic and smooth, can be expressed as the sum of sinusoidal harmonic components.

Beats

Beats occur when two sinusoidal oscillations, whose frequencies do not differ greatly, are superposed ($f_1 \approx f_2$). They are periodic. Their basic frequency is the difference between the frequencies of the superposed sinusoidal oscillations.

Natural frequency

The natural frequency is that frequency f at which an oscillating system can oscillate freely after being excited once (natural oscillation). It is dependent only on the properties of the oscillating system.

Damping

Damping is a measure of the energy losses in an oscillatory system when one form of energy is converted into another. The consequence is a decay of the oscillation (Figure 2).

Damping ratio

The damping ratio D is the measure for the degree of damping.

Logarithmic decrement

The logarithmic decrement Δ is the natural logarithm of the relationship between two extreme values of a damped natural oscillation which are separated by one period.

Forced oscillations

Forced oscillations arise under the influence of an external physical force (excitation), which does not change the properties of the oscillator. The frequency of forced oscillations is determined by the frequency of the excitation.

Transfer function

The transfer function is the quotient of amplitude of the observed state or output variable and the amplitude of excitation, plotted against the excitation frequency f or the excitation-circuit frequency ω .

Resonance

A resonance occurs when the transfer function attains its maximum value as the excitation frequency approaches the natural frequency.

Resonant frequency

The resonant frequency is the excitation frequency at which the oscillator state variable attains its maximum value. Disregarding the damping, the resonant frequency is equal to the natural frequency.

Half-value width

The half-value width is the difference between the frequencies at which the level of the observed variable has dropped to $1/\sqrt{2} \approx 0.707$ of the maximum value.

Resonance sharpness

The resonance sharpness Q (or quality factor) is the maximum value of the transfer function.

Coupling

If two oscillatory systems are coupled together – mechanically by mass or elasticity, electrically by inductance or capacitance – a periodic exchange of energy takes place between the systems.

Wave

A wave is a spatial and temporal change of state of a continuum, which can be expressed as a unidirectional transfer of location of a certain state over a period of time. The matter that may be present in the space is not necessarily also transported.

There are transversal waves (e.g. waves in rope and water) and longitudinal waves (e.g. sound waves in air).

Interference

The principle of undisturbed superposition of waves is called interference. At every point in space, the instantaneous value of the resulting wave is equal to the sum of the instantaneous values of the individual waves.

Plane wave

A plane wave is a wave in which the surfaces of the same phase (e.g. maxima or wave fronts) form a plane, i.e. the wave propagates linearly. The wave fronts are vertical to the direction of propagation.

Standing waves

Standing waves occur as a result of interference between two waves of equal frequency, wavelength, and amplitude traveling in opposite directions. In contrast to a propagating wave, the amplitude of the standing wave is constant at every point; nodes (zero amplitude) and antinodes (maximum amplitude) occur. Standing waves occur, for example, when a plane wave is reflected on a plane wall which is vertical to the direction of wave propagation.

Rectification value

The rectification value y_{rec} is the arithmetic mean value, linear in time, of the values of a periodic signal:

$$y_{\text{rec}} = \frac{1}{T} \int_0^T |y| dt.$$

For a sine curve:

$$y_{\text{rec}} = \frac{2\hat{y}}{\pi} \approx 0.637 \hat{y}.$$

Effective value

The effective value y_{eff} is the time virtual value of a periodic signal. It is also known as the RMS value (root mean square):

$$y_{\text{eff}} = \sqrt{\frac{1}{T} \int_0^T y^2 dt}.$$

For a sine curve:

$$y_{\text{eff}} = \frac{\hat{y}}{\sqrt{2}} \approx 0.707 \hat{y}.$$

Harmonic factor

The harmonic factor is the ratio of y_{eff} to y_{rec} . For a sine curve the harmonic factor is $y_{\text{eff}}/y_{\text{rec}} \approx 1.111$.

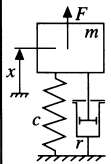
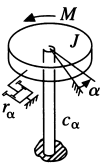
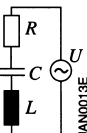
Peak factor

For a sine curve the peak factor is $\hat{y}/y_{\text{eff}} = \sqrt{2} \approx 1.414$.

Equations

The following equations apply to simple oscillators (Table 2) if the general quantity designations in the formulas are replaced by the relevant physical quantities.

Table 2: Simple oscillatory systems

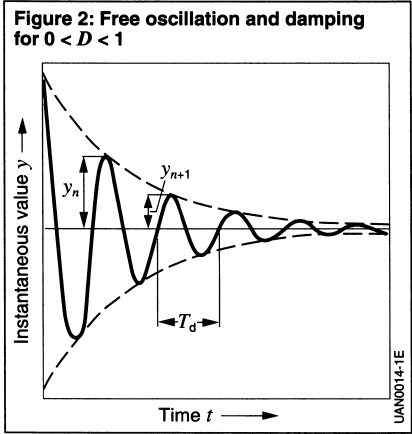
	Mechanical		Electrical
	Translational	Rotational	
			
Des.	Physical quantity		
y	x	α	Q
$\dot{y}$	$\dot{x} = v$	$\dot{\alpha} = \omega$	$\dot{Q} = I$
$\ddot{y}$	$\ddot{x} = \dot{v}$	$\ddot{\alpha} = \dot{\omega}$	$\ddot{Q} = \dot{I}$
F_Q	F	M	U
a	m	J	L
b	r	r_α	R
c	c	c_α	$1/C$

Differential equation

$$a\ddot{y} + b\dot{y} + cy = F_Q(t) = \hat{F}_Q \sin \Omega t,$$

Period $T = 1/f$,
Angular frequency $\omega = 2\pi f$.

Sinusoidal oscillation: $y = \hat{y} \sin \omega t$.



Free oscillations ($F_Q = 0$)

Logarithmic decrement (Figure 2):

$$\Lambda = \ln \left(\frac{y_n}{y_{n+1}} \right) = \frac{\pi b}{\sqrt{ca - \frac{b^2}{4}}}$$

$$\text{Decay coefficient } \delta = \frac{b}{2a},$$

$$\text{Damping ratio } D = \frac{\delta}{\omega_0} = \frac{b}{2\sqrt{ca}},$$

$$D = \frac{\Lambda}{\sqrt{\Lambda^2 + 4\pi^2}} \approx \frac{\Lambda}{2\pi} \quad (\text{low level of damping}).$$

Angular frequency of undamped oscillation ($D = 0$): $\omega_0 = \sqrt{c/a}$.

Angular frequency of damped oscillation ($0 < D < 1$): $\omega_d = \omega_0 \sqrt{1 - D^2}$.

For $D \geq 1$ no oscillations, but creepage.

Forced oscillations

Quantity of transfer function:

$$\frac{\hat{y}}{\hat{F}_Q} = \frac{1}{\sqrt{(c - a\Omega^2)^2 + (b\Omega)^2}}$$

$$= \frac{1}{c \sqrt{\left(1 - \left(\frac{\Omega}{\omega_0}\right)^2\right)^2 + \left(\frac{2D\Omega}{\omega_0}\right)^2}},$$

Resonant frequency $f_g = f_0 \sqrt{1 - 2D^2} < f_0$,

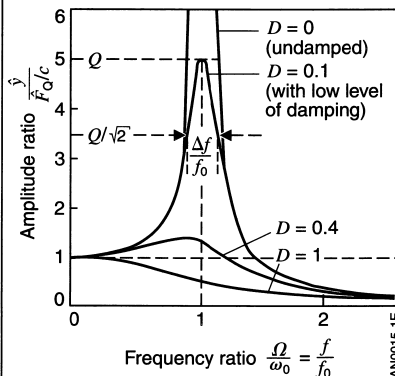
Resonance sharpness $Q = 1/(2D\sqrt{1 - D^2})$,

Resonant frequency $f_g \approx f_0$ (for $D \leq 0.1$),

Resonance sharpness $Q \approx \frac{1}{(2D)}$ (for $D \leq 0.1$),

Half-value width $\Delta f = 2Df_0 = \frac{f_0}{Q}$.

Figure 3: Standardized transfer function



Vibration reduction

Vibration damping

If damping can only be carried out between the machine and a quiescent point, damping must be at a high level (Figure 3).

Vibration isolation

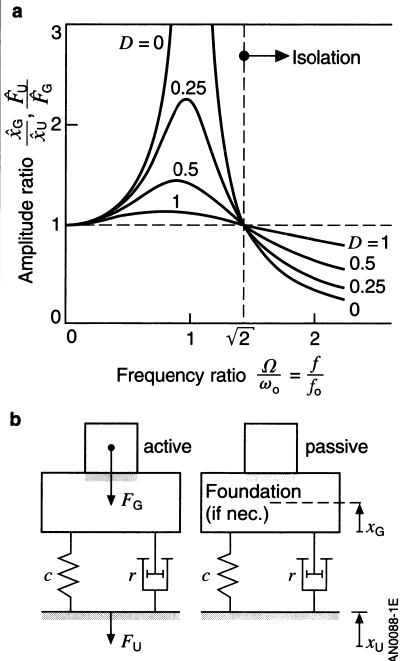
Active vibration isolation

The machines to be isolated are to be mounted so that the dynamic forces transmitted to the base support are small.

One measure to be taken: The bearing point should be set below resonance so that the natural frequency lies below the lowest excitation frequency. Damping impedes insulation. Low values can result in excessively high vibrations during run-up when the resonant range is passed through (Figure 4).

Figure 4: Vibration isolation

a) Transfer function,
b) Low setting.



Passive vibration isolation

The machines to be isolated are to be mounted so that vibrations and shocks reaching the base support are only transmitted to the machines to a minor extent.

The measures to be taken are the same as those for active isolation. In many cases, flexible suspension or extreme damping is not practicable. To prevent the occurrence of resonance, the machine attachment should be so rigid that the natural frequency is far enough in excess of the highest excitation frequency which can occur (Figure 4).

Vibration absorption

Vibration absorber with fixed natural frequency

By tuning the natural frequency f_T of an absorption mass with a flexible, loss-free coupling to the excitation frequency, vibrations acting on the machine are completely absorbed (Figure 5). Only the absorption mass still vibrates. The effectiveness of the absorption decreases as the excitation frequency changes. Damping prevents complete absorption. How-

ever, appropriate tuning of the absorber frequency and an optimum damping ratio produce broadband vibration reduction, which remains effective when the excitation frequency changes.

Vibration absorber

with varying natural frequency

Rotational oscillations with excitation frequencies proportional to the rotational speed (e.g. orders of balancing in IC engines) can be absorbed by absorbers with natural frequencies proportional to the rotational speed (pendulum in the centrifugal-force field). The vibration absorption is effective at all rotational speeds. Vibration absorption is also possible for oscillators with several degrees of freedom and interrelationships, as well as by the use of several absorption masses.

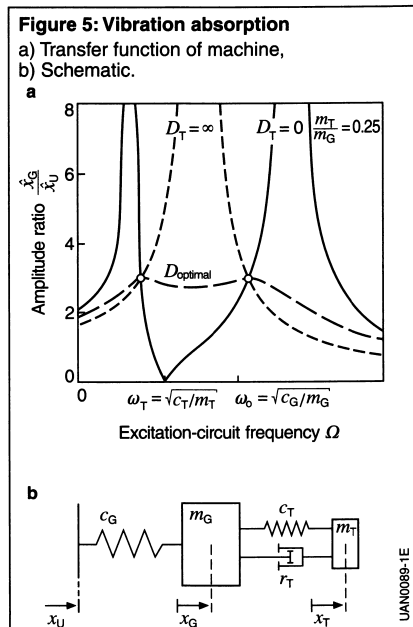
Modal analysis

The dynamic behavior (natural-oscillation characteristics) of oscillating systems is determined by means of modal analysis. It is used among other things in design to optimize structures with regard to the oscillatory characteristics and to identify problem areas, and in acoustics to analyze structure-borne noise ([3]).

The oscillating structure, which as a continuum has infinitely many degrees of freedom, is replaced in a clearly defined manner by a finite number of single-mass oscillators. The modal model of the structure created in this way is described by the modal parameters

- natural-oscillation shapes (also eigenvector or mode),
- natural frequencies (also eigenvalues),
- and the associated modal damping values.

A time-invariant and linear-elastic structure is an essential precondition for model creation. Every oscillation of the structure can be represented from the eigenvectors and eigenvalues. It is, however, only observed at a limited number of points in the possible oscillation directions (degrees of freedom) and at defined frequency intervals.



The substructure-coupling process collates modal models of various structures into an overall model.

Numerical modal analysis

The geometry, material data, and marginal conditions must be known. The basis for numerical modal analysis is a multi-body-system or finite-element model of the structure. Eigenvalues and eigenvectors can be calculated from this by solving an eigenvalue problem.

Numerical modal analysis manages without a prototype of the structure and can already be used at an early stage of development. However, it is often the case that precise knowledge concerning the structure's fundamental properties (damping, marginal conditions) are lacking, which means that the modal model can sometimes be inaccurate. As well as this, the error is unidentified. One remedy is to adjust the model with the results of an experimental modal analysis.

Experimental modal analysis

A prototype of the structure is required for experimental modal analysis. Analysis is based on measurements of transfer functions. To this end, either the structure is excited at one point in the frequency range of interest and the oscillation responses measured at several points, or it is excited at many points in succession, where the oscillation responses are always measured at the same point. The modal model is derived from the matrix of the transfer function (it describes the response model). An impulse hammer or an electrodynamic or hydraulic "shaker" is used as the means of excitation. The response is measured with acceleration sensors or a laser vibrometer.

Experimental modal analysis can also be used to validate numerical analysis. Simulation calculations can then be carried out on the validated numerical model. In the response calculation, the response of the structure to a defined excitation is calculated, an excitation which corresponds, for example, to test-bay conditions.

By means of structure modifications (changes in mass, damping or stiffness), the vibrational behavior can be opti-

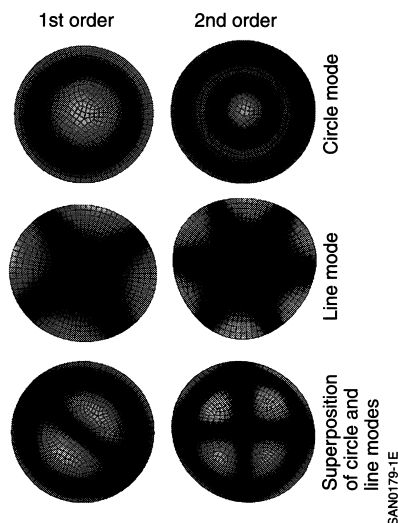
mized to the level required by operating conditions. When the modal models produced by both processes are compared with each other, the modal model resulting from an analytical modal analysis is more detailed than that from an experimental modal analysis, due to the greater number of degrees of freedom in the analytical process. This applies in particular to simulation calculations based on the model.

The natural-oscillation shapes resulting from a modal analysis can be shown in graphic form or animated (examples in Figure 6). The different gray stages describe the excursion vertically to the plane of projection. The partial distortion of the disks results from these excursions.

References

- [1] DIN 1311: (Mechanical) vibrations, oscillation and vibration systems – Part 1: Basic concepts, survey.
- [2] P. Hagedorn, D. Hochlenert: Technische Schwingungslehre. 2nd Ed., Europa Lehrmittel, 2014.
- [3] H.G. Natke: Einführung in Theorie und Praxis der Zeitreihen- und Modalanalyse. 3rd Ed., Vieweg+Teubner Verlag, 2013.

Figure 6: Eigenmode of round disks



Acoustics

Acoustics (Greek “to hear”) is a branch of the general science of vibration and oscillation concerned with the vibration of air particles audible to humans. This process of vibration gives rise to pressure differences in air which the human ear can detect. Vibrations are generated on an automobile by a variety of excitations which partly excite the airborne noise directly or can propagate via the vehicle structure and from there to the airborne noise.

General terminology

(Symbols and units in Table 1, see also DIN 1320 [1])

Sound

Mechanical vibrations and waves in an elastic medium in the audible frequency band (16 to 20,000 Hz) are called sound.

Ultrasound

High-frequency vibrations above the frequency range of human hearing are called ultrasound.

Infrasound

Infrasound is the range of very low frequencies which can no longer be detected by the human hear, but can still be felt by the body.

Sound pressure

Sound pressure p is the alternating pressure generated in a medium by the vibration of sound.

Particle velocity

Particle velocity v is the alternating velocity of a vibrating particle. In a free sound field:

$$v = \frac{p}{Z}.$$

Specific acoustic impedance

Specific acoustic impedance Z denotes the wave impedance of a medium and is a measure of the ability of a medium to transmit sound waves.

In a free field the relationship between sound pressure, particle velocity and acoustic impedance is:

$$Z = \frac{p}{v} = \rho c.$$

Figuratively speaking, the formula describes the coupling of a vibrating particle to its neighboring particles. The coupling is much less pronounced in gases than in solid media; the particles can vibrate more freely.

Table 1: Quantities and units
(see also DIN EN ISO 80000-8 [2])

Quantity	SI unit
c Velocity of sound	m/s
f Frequency	Hz
I Sound intensity	W/m ²
L_I Sound intensity level	dB
L_{Aeq} Equivalent continuous sound level, A-weighted	dB (A)
L_{den} Noise indices	dB (A)
L_{pA} Sound pressure level, A-weighted	dB (A)
L_r Rating sound level	dB (A)
L_{WA} Sound power level, A-weighted	dB (A)
L_S Loudness level	phon
S Loudness	sone
P Sound power	W
p Sound pressure	Pa
S Surface area	m ²
v Particle velocity	m/s
Z Acoustic impedance	Pa · s/m
α Sound absorption coefficient	1
λ Wavelength	m
ρ Density	kg/m ³
ω Angular frequency (= $2 \pi f$)	1/s
SEL Sound exposure level	dB (A)

Hearing dynamics

Hearing dynamics is the ability of the human ear to perceive noises from the threshold of audibility up to the pain threshold. It corresponds to pressure differences from the threshold of audibility, which is generally defined as 20 μPa , up to the pain threshold of approx. 200 pascal. This results in very high hearing dynamics of around 1:10,000,000.

Decibel dB

The decibel is not a unit, but the logarithmic ratio of a measured quantity to a reference value. The use of decibels makes it easier to handle the large dynamic range.

The sound pressure level is given in dB referred to the reference pressure of the threshold of audibility ($p_{\text{ref}} = 20 \mu\text{Pa}$):

$$L_p = 20 \log_{10} \left(\frac{p}{p_{\text{ref}}} \right).$$

Frequency

Frequency f specifies the number of vibrations per second. The unit is the hertz (Hz) and is defined as follows:

$$1 \text{ Hz} = 1/\text{s}.$$

Velocity of sound

Velocity of sound c is the velocity of propagation of a sound wave in a medium (Table 2).

Wavelength

The distance between two wave crests is called the wavelength. It is linked to the frequency:

$$\lambda = \frac{c}{f} = \frac{2\pi c}{\omega}.$$

Propagation of sound

In general, sound propagates spherically from its source. In a free sound field, the sound pressure level decreases by 6 dB each time the distance from the sound source is doubled ($20 \log 2$). Reflecting objects influence the sound field, and the rate at which the sound level is reduced as a function of the distance from the sound source is lower. Moving vehicles have a spherical propagation of sound, but are considered to be line sources. The damping of sound each time the distance is doubled is considerably less and ranges, depending on the vehicle and the computing model, between 3 dB and 4.5 dB each time the distance is doubled.

Sound power

Sound power P is the sound energy output by a sound source per unit of time. It is not dependent on space and distance.

Sound intensity

Sound intensity I is the sound power through a plane vertical to the direction of propagation:

$$I = \frac{P}{S}.$$

In a sound field:

$$I = \frac{p^2}{\rho} c = v^2 \rho c.$$

Doppler effect

For moving sound sources: As the sound source approaches the observer, the perceived pitch is higher than the actual pitch; as distance increases, the perceived pitch falls.

Table 2: Sound velocities and wavelengths in different materials

Material/medium	Sound velocity c in m/s	Wave-length λ in m at 1,000 Hz
Air, 20 °C, 1,014 hPa	343	0.343
Water, 10 °C	1,440	1.44
Rubber (acc. to hardness)	60–1,500	0.06–1.5
Aluminum (rod)	5,100	5.1
Steel (rod)	5,000	5.0

Sound spectrum

Every noise is a mixture of components with different frequencies and levels. The sound-pressure level components can be broken down by frequencies with a frequency analysis. These spectra are distinguished from each other, depending on the type of frequency resolution.

Octave band spectrum

The sound pressure levels are determined and represented in terms of octave bandwidth. In the case of the octave, the fundamental frequencies behave in a ratio (f_1, f_2) of 1:2. The mean frequency of the octave is:

$$f_m = \sqrt{f_1 f_2}.$$

Third-octave band spectrum

The sound pressure levels are determined and represented in terms of third-octave bandwidth. In the case of the third, the fundamental frequencies behave in a ratio of $1:2^{1/3}$. The bandwidth referred to the center frequency is relatively constant, as in the case of the octave band spectrum.

The frequencies are defined in DIN EN 61260 [3]. Frequency analyses in the octave and third-octave ranges produce bar charts with the frequency on the x -axis and the associated levels on the y -axis.

Narrow-band spectrum

Unlike the above spectra, it is possible to use Fourier analysis [4] to analyze frequency components with constant frequency bandwidths. This results in very many finer frequency resolutions compared with the above spectra. These are shown as line graphs, which strictly speaking are not correct, but enable different analyses to be compared.

Such narrow-band analyses are also frequently linked with a further piece of information such as, for example, the rotational speed of an engine so that a frequency analysis is available for each rotational speed (for example when the engine is revved up). Such analyses are shown in three-dimensional form, either as "waterfall charts" or today frequently in color spectra, with the x -axis showing the rota-

tional speed, the y -axis the frequency, and the z -axis in its aspect the color coding.

Sound insulation

Sound insulation is achieved by interposing a reflecting (insulating) wall (e.g. a building wall) between the sound source and the impact location. This reduces the effect of the sound source.

Sound damping, sound absorption

In the case of sound damping and sound absorption, sound energy can penetrate into the medium. Here, energy is converted into heat when reflected on peripheries, but also during propagation in the medium.

Modern noise baffles work on the principle of sound damping, i.e. sound is reflected. But parts of the sound energy are also absorbed by the baffles.

Sound absorption coefficient

The sound absorption coefficient α is the ratio of non-reflected to incident sound energy. With total reflection, $\alpha = 0$; with total absorption, $\alpha = 1$.

Low-noise design

A low-noise design is a design which is structurally optimized in accordance with acoustic considerations in order to minimize the inherent sound radiation and the propagation of sound within the structure. Simulation techniques for calculating and optimizing acoustic properties are frequently used.

Noise reduction

Noise reduction involves reducing the noise emissions of a complete system, primarily by means of low-noise designs and, secondly, through the reduction of sound propagation through the use of insulating, damping and absorbing materials.

Measured quantities for noise emissions

Sound-field quantities are usually given as effective values. Because humans perceive frequencies with varying loudness, weighting filters are used to adapt the measured levels in each frequency range to the human hearing properties. The most common weightings are A for sound as encountered, for example, in the automobile sector, and C for clearly louder noises such as, for example, in the aviation field. The corresponding weighting is appended to the symbol, such as, for example, dB(A).

Sound power level

The sound power of a sound source is described by the sound power level L_w . It is equal to ten times the logarithm to the base 10 of the ratio of the calculated sound power to the reference sound power $P_0 = 10^{-12}$ W:

$$L_w = 10 \log \left(\frac{P}{P_0} \right),$$

where log here and also in the following means the logarithm to the base 10.

Sound power cannot be measured directly. It is calculated based on quantities of the sound field which surrounds the source. The sound pressure levels L_p or the sound intensity levels L_I are usually used for the calculation.

Sound pressure level

The sound pressure level L_p is ten times the logarithm to the base 10 of the ratio of the square of the effective sound-pressure value and the square of the reference sound pressure $p_0 = 20 \mu\text{Pa}$:

$$L_p = 10 \log \left(\frac{p_{\text{eff}}^2}{p_0^2} \right) \text{ or}$$

$$L_p = 20 \log \left(\frac{p_{\text{eff}}}{p_0} \right).$$

The sound pressure level is given in decibels (dB). The frequency-dependent, A-weighted sound pressure level L_{pA} as measured at a distance of $d = 1$ m is frequently used to characterize sound sources.

Sound intensity level

The sound intensity level L_I is ten times the logarithm to the base 10 of the ratio of sound intensity and reference sound intensity $I_0 = 10^{-12}$ W/m²:

$$L_I = 10 \log \left(\frac{I}{I_0} \right).$$

The sound intensity can be measured directly with a probe.

Interaction of two or more sound sources

If two independent sound fields are superimposed, their sound intensities or the squares of their sound pressures must be added. The overall sound level is then determined from the individual sound levels in accordance with Table 3.

Table 3: Overall sound level with superimposition of independent sound fields

Difference between 2 individual sound levels	Overall sound level = higher individual sound level + allowance of:
0 dB	3.0 dB
1 dB	2.5 dB
2 dB	2.1 dB
3 dB	1.8 dB
4 dB	1.5 dB
6 dB	1.0 dB
8 dB	0.6 dB
10 dB	0.4 dB

Measured quantities for noise immissions

Excessive noise

Excessive noise is the classification of a noise as an undesirable sound event and is dependent on the following factors:

- on the noise itself, measurable by physical variables (e.g. frequency, sound pressure level, or sound power),
- on the subjective attitude to the noise of the person affected,
- on the personal state of the person affected, and
- on the concrete situation in which the noise occurs.

Protection against excessive noise and in particular environmental noise is gaining increasing importance as an environmental theme. The recording and evaluation of noise immissions forms the basis for efficient noise reduction.

Rating sound level

The effect of noise on a human being is evaluated using the rating sound level L (see also DIN 45645-1 [5]; an ISO standard is currently being worked on). This is a measure of the mean noise immission over a period of time (e.g. eight working hours). With fluctuating noises it is either measured directly with integrated measuring instruments or calculated from individual sound-pressure-level measurements and the associated periods of time of the individual sound effects (see also DIN 45641, [6]). Noise-immission parameters such as pulsation (short strong deviations of the sound level from the mean level) and tonal quality (heavy dominance of one or more discrete frequencies) can be taken into account through level allowances.

Energy-equivalent continuous sound level

In the case of noises which fluctuate over time, the mean A-weighted sound pressure level resulting from the individual sound pressure levels and the individual exposure times equals the energy-equivalent continuous sound level L_{Aeq} (see also DIN 45641).

Table 4 provides guide values for the rating sound level (Germany; TI Noise [7]), measured outside the nearest residential building (0.5 m in front of an open window).

Noise indices

The EU Directive 2002/49/EC [8] defines the noise indices binding on the EU L_{den} and L_{night} as standard descriptors for offending noise over a full day (L_{den} for day, evening, night) and a night (L_{night}) along similar lines to the above-described L_{Aeq} . The assessment period here is one calendar year. Crashes for evening (5 dB) and night (10 dB) are taken into consideration for L_{den} in contrast to L_{Aeq} .

$$L_{den} = 10 \log \frac{1}{24} \cdot \left(12 \cdot 10^{\frac{L_d}{10}} + 4 \cdot 10^{\frac{L_e + 5}{10}} + 8 \cdot 10^{\frac{L_n + 10}{10}} \right)$$

where L_d , L_e and L_n are the values for day, for evening, and for night.

Sound exposure level

The sound exposure level SEL is used to evaluate individual excessive-noise events (e.g. an airplane starting) in sensitive conservation areas. For the purpose of calculation the sound energy of the complete event is recorded and its energy is then distributed to 1 second. It is therefore done as though the event always happens within a period of one second. The sound level is then calculated on this basis.

Table 4: Guide values for permissible noise pollution as per TI Noise [7]

	Day	Night
Purely industrial areas	70 dB (A)	70 dB (A)
Areas with predominantly commercial/industrial premises	65 dB (A)	50 dB (A)
Mixed areas	60 dB (A)	45 dB (A)
Areas with predominantly residential premises	55 dB (A)	40 dB (A)
Purely residential areas	50 dB (A)	35 dB (A)
Health resorts, hospitals, etc.	45 dB (A)	35 dB (A)

Perceived sound levels

The human ear can distinguish approximately 300 levels of acoustic intensity and 3,000 to 4,000 different frequencies (pitch levels) in rapid temporal succession and evaluate them according to complex patterns. However, there is not necessarily any direct correlation between perceived loudness levels and (energy-oriented) technically defined sound levels. Corrections (A-, B- and C-weighting) take into account the dependence on frequency of human hearing (Figure 1).

The phon unit and the definition of loudness in "sone" provide a rough approximation of subjective sound-level perception. Sound-level measurements alone do not suffice to define the nuisance and disturbance potential of noise emanating from machinery and equipment. A hardly perceptible ticking noise can thus be perceived as extremely disturbing, even in an otherwise loud environment.

Loudness level

The loudness level L_S is a comparative measure of the intensity of the subjective perception of a sound, measured in "phon". The loudness level of a sound (pure tone or noise) is the sound pressure level of a standard pure tone which, under standard listening conditions, is judged by a normal observer to be equally loud. The standard noise level is a plane sound wave at a frequency of 1,000 Hz impinging on the observer's head from the front.

This is known internationally as the "loudness level". A difference of 8 to 10 phon is perceived as twice or half as loud.

Phon

The standard pure tone judged as being equally loud has a specific value in dB. This value is given as the loudness level of the tested sound, and has the designation "phon". As human perception of sound is frequency-dependent, the dB values of the tested sound for notes, for example, do not agree with the dB values of the standard pure tone. The connection between curves of equal sound-level perception and sound pressure in decibels is always empirically determined. These curves were published for the first time in 1933 by Fletcher and Munson. They provided the basis for the isophone curves in accordance with DIN ISO 226 [9] used today. Figure 2 shows curves based on this standard.

Loudness in "sone"

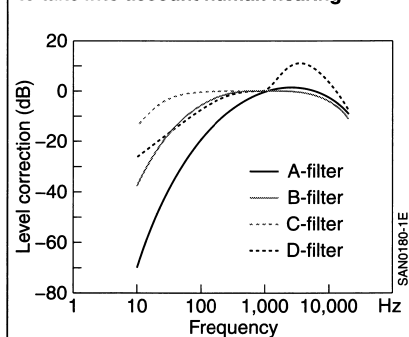
The loudness S is the measure used to define subjective noise levels. The starting point for defining the sone is: How much higher or lower is the perceived level of a particular sound relative to a specific standard. Definition: Loudness level of $L_S = 40$ phon corresponds to a loudness $S = 1$ sone. Doubling or halving the loudness is equivalent to a variation in the loudness level of approx. 10 phon.

There is a DIN loudness standard for calculating stationary sound using tertiary levels (Zwicker method [10]). This procedure takes into account both frequency weighting and the screening effects of hearing.

Pitch, sharpness

The spectrum of perceptible sound can be divided into 24 hearing-oriented frequency groups (bark). The Bark scale is a psychoacoustic scale for perceived pitch. The loudness/pitch distribution can be used to quantify other subjective aural impressions – such as the sharpness (e.g. metallic harmonic distortion) of a noise.

Figure 1: Frequency-dependent correction values for the sound level to take into account human hearing



Articulation index

The basic prerequisite for recognizing speech is the correct transmission wherever possible of the sound waves of what is spoken from the sender (mouth) to the receiver (ear). The articulation index (AI) offers a computing model for predicting speech intelligibility in the event of interference noise. It is assumed here that the speech information is distributed to the different frequency bands of the acoustic signal. Each band then makes a contribution to speech intelligibility. If a band has a signal-to-noise ratio of more than 15 dB, the speech component in this band is evaluated as being intelligible. The sum total of all the bands with this noise ratio is multiplied by their specific weighting factor to give the articulation index AI in %.

References

- [1] DIN 1320: Acoustics – Terminology.
- [2] DIN EN ISO 80000-8: Quantities and units – Part 8: Acoustics.
- [3] DIN EN 61260: Electroacoustics – Octave-band and fractional-octave-band filters.
- [4] B. Lenze: Einführung in die Fourier-Analyse. 3rd Ed., Logos Verlag, Berlin, 2000.
- [5] DIN 45645-1: Determination of rating levels from measurement data – Part 1: Noise immission in the neighbourhood.
- [6] DIN 45641: Averaging of sound levels.
- [7] Sixth General Administrative Regulation on the Federal Immission Protection Law (Technical Instructions on Noise Abatement – TI Noise) of 26 August 1998 (GMBI No. 26/1998 P. 503).
- [8] Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise.
- [9] DIN ISO 226: Acoustics – Normal equal-loudness-level contours.
- [10] DIN 45631: Calculation of loudness level and loudness from the sound spectrum – Zwicker method.

Figure 2: Isophone curves in acc. with DIN ISO 226 with an allocation of noises (examples)

Sensation of pain
4-engine airplane
(3 meters away)

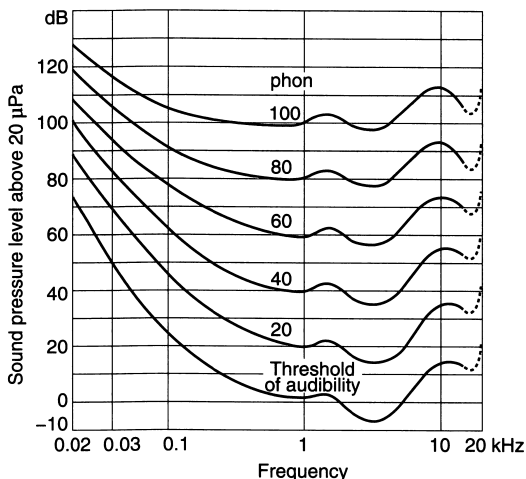
Boilermaking

Pneumatic hammer

Truck, bus
Passenger car, loud speech
(1 meter away)

Normal conversation

Radio at low volume
Living room by day
Living room at night
Ticking of a clock
Rustling of leaves
Threshold of audibility



Optical technology

Electromagnetic radiation

Electromagnetic radiation denotes the wave nature of visible light, but also of radiation that is not visible to the human eye (e.g. ultraviolet light, X-radiation, radiation in the infrared range). Thus, electromagnetic radiation is subdivided by reference to its wavelength λ (Table 2).

Table 1: Quantities and units

Quantity	SI unit
A Area	m^2
A_1 Radiating area	
A_2 Illuminated area	
E_e Irradiance	W/m^2
I_e Radiant intensity	W/sr
L_e Radiance	$\text{W}/(\text{m}^2 \text{sr})$
M_e Radiant excitance	W/m^2
P_e Power	W
Q_e Radiant energy	Ws
R Reflection coefficient	$\%$
H_e Irradiation	Ws/m^2
$K(\lambda)$ Absolute spectral sensitivity	lm/W
$V(\lambda)$ Spectral luminous efficiency	—
Φ_e Radiant flux	W
r Distance	m
t Time	s
e_1 Angle, incident ray (to normal of a surface)	$^\circ$
e_2 Angle, refracted ray	$^\circ$
e_3 Angle, reflected ray	$^\circ$
$e_{1,\text{max}}$ Angle of total reflection	$^\circ$
η Luminous efficiency	lm/W
Ω Solid angle	sr
λ Wavelength	nm
E_{phot} Photon energy	eV
h Planck's quantum of action	Js
c Speed of light in a vacuum	m/s
n Refractive index	—

Geometric optics

The description of optical systems can be limited to two models – geometric optics and wave optics. Geometric optics (also called ray model) describes the propagation of light and its deflection by imaging elements with simple geometric considerations. But this is only permitted provided the imaging elements (e.g. mirrors, lenses) are much larger than the wavelength of light.

Effects such as interference, diffraction and polarization are described in physics by wave optics. This is concerned with the propagation of light in the form of a wave.

Basic principles of geometric optics

A large part of the imaging systems used in optics can be described by geometric optics. Radiation propagation is explained by “light beams” and can be described by means of simple geometric laws.

Table 2: Ranges of electromagnetic radiation

Designation	Wavelength range
Gamma radiation	0.1...10 pm
X-radiation	10 pm to 10 nm
Ultraviolet radiation	10...380 nm
Visible radiation	380...780 nm
Infrared radiation	780 nm to 1 mm
Millimeter waves (EHF)	1...10 mm
Centimeter waves (SHF)	10...100 mm
Decimeter waves (UHF)	100 mm to 1 m
Ultrashort waves (VHF)	1...10 m
High-frequency waves (HF)	10...100 m
Medium waves (MF)	100 m to 1 km
Long waves (LF)	1...10 km
Myriameter waves (VLF)	10...100 km

Basic axioms of geometric optics

Fermat's principle is the basic principle for geometric optics. The basic axioms of geometric optics can be derived from the basic statement of this principle to the effect that light always chooses the shortest path between two points [1]:

- Light beams are straight in homogeneous media.
- At the boundary between two homogeneous materials the beams are deflected according to the law of reflection and the law of refraction.
- Each beam path is reversible.
- Light beams can cross without influencing each other.

Law of refraction

Snell's law of refraction [2] can also be derived from Fermat's principle. This law describes the transition of a light beam from one medium (e.g. air) to another medium (e.g. glass). At the interface of the two media the light is deflected due to the change of refractive index. An incident ray is split into a refracted ray and a reflecting ray (Figure 1). The refracted ray is governed by the law of refraction:

$$n_1 \sin \varepsilon_1 = n_2 \sin \varepsilon_2.$$

For vacuum and dielectric media (e.g. air, glass, plastics) the refractive indices n are real numbers (Table 3). Materials with a large refractive index n are referred to as optically thick, materials with a small refractive index as optically thin.

The characteristic of refractive indices of being dependent on the wavelength is called dispersion. In most cases they decrease as the wavelength increases.

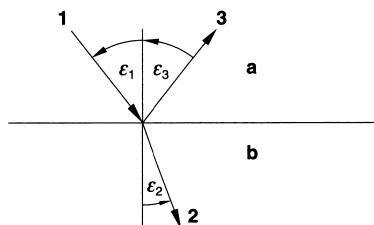
Law of reflection

The following formula applies to the direction of the reflecting ray (law of reflection):

$$\varepsilon_3 = \varepsilon_1.$$

Figure 1: Refraction of a light beam at the interface between two media with different refractive indices

- a Medium 1 with refractive index n_1 ,
 b Medium 2 with refractive index n_2 .
 1 Incident ray,
 2 Refracted ray,
 3 Reflected ray.
 ε_1 Angle, incident ray,
 ε_2 Angle, refracted ray,
 ε_3 Angle, reflected ray.



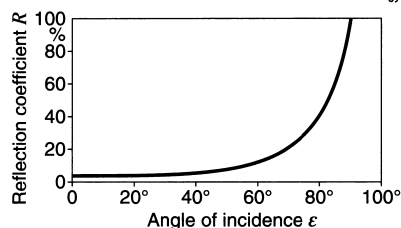
UAN0028-1Y

Table 3: Refractive index n of some media
 (For yellow light of wavelength $\lambda = 589.3$ nm)

Vacuum, air	1.00
Ice (0°C)	1.31
Water (20°C)	1.33
Quartz glass	1.46
Polymethylmethacrylate	1.49
Standard glass for optics (BK 7)	1.51673
Window glass	1.52
Glass for headlamp lenses	1.52
Polyvinylchloride	1.54
Polycarbonate	1.58
Polystyrene	1.59
Epoxy resin	1.60
Gallium arsenide (depending on doping)	approx. 3.5

Figure 2: Dependence of the reflection coefficient on the angle of incidence

Medium 1 with refractive index $n_1 = 1.00$,
 Medium 2 with refractive index $n_2 = 1.52$.



SAN0198E

The angle of reflection is therefore equal to the angle of incidence. The intensity ratio of reflected to incident radiation (degree of reflection, reflection coefficient) is described by the Fresnel formula [3] and is dependent on the angle of incidence ε_1 and on the refractive indices of the adjoining media (Figure 2). In the transition from air ($n_1 = 1.00$) to glass ($n_2 = 1.52$) the upshot is that, in the case of vertical ray incidence ($\varepsilon_1 = 0$), 4.3% of the ray energy is reflected.

Total reflection

In the transition from an optically thicker to an optically thinner medium ($n_1 > n_2$) the angle of refraction ε_2 attains for a specific angle of incidence ε_1 its highest possible value (90°). This angle of incidence $\varepsilon_{1,\max}$ is called the critical angle for total reflection. According to the law of refraction, the following is obtained for total reflection:

$$\sin \varepsilon_{1,\max} = \sin 90^\circ \frac{n_2}{n_1} = \frac{n_2}{n_1}.$$

Optical components

Spherical lenses

The optical effect and the imaging property of a lens are determined by the shape of the interfaces and by the refractive index of the lens material [4]. The imaging properties of light beams can be determined using Snell's law of refraction.

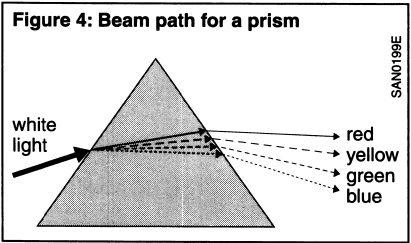
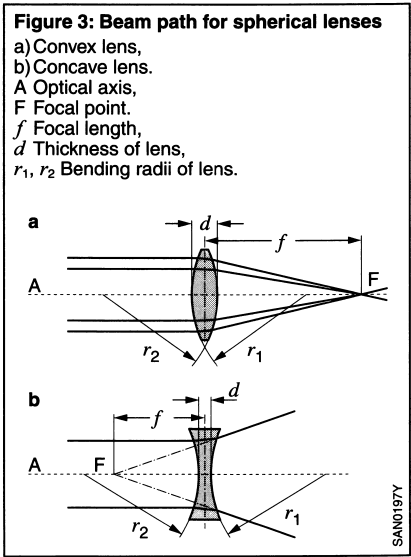
The functioning principle of a lens is divided into two subgroups based on its shape (Figure 3). Lenses that bring light to a focal point are called convergent lenses (convex lenses). In contrast, parallel beams of rays are spread out by divergent lenses (concave lenses).

Prisms

In the case of a prism, the dependence of the refractive index on the wavelength of the incident light is utilized (dispersion). Thus, blue light ($\lambda = 420$ to 490 nm) is refracted at a different angle ε_2 from red light ($\lambda = 650$ to 750 nm). When white light shines onto a prism, it is broken down into different component parts (color components) (Figure 4).

This effect also appears in rainbows. The different color components of sunlight are refracted differently on water droplets in the air.

Vehicle headlights use lenses (elements comprising cylindrical lenses and prisms) to favorably influence the radiation coming from the reflector.



Reflectors

Reflectors (mirrors) are used in optics to divert light in a particular direction (Figure 5). Reflectors are used in among other things headlights, where they specifically direct and concentrate light.

The function of reflector is to capture light from the headlight bulb, to achieve as great a range as possible, and to influence the distribution of the light on the road in such a way as to satisfy the legal requirements. Design (e.g. when installed in the bumper) places additional demands on the headlights.

Whereas previously almost exclusively paraboloids were used for reflectors, the above-mentioned, partly contradictory requirements can today sometimes only be satisfied by stepped reflectors, free-form surfaces or new headlight concepts (see Lighting equipment).

Basically speaking, the larger the lens aperture area, the greater the headlight range that can be achieved. On the other hand, the luminous efficiency increases with the size of the solid angle captured by the reflector.

Color filters

For special applications, e.g. lights and lamps on vehicles, there are precise regulations regarding the wavelength of the light used to suit the specific purpose (turn-signal lamp, stop lamp). Color filters are used here to attenuate or suppress unwanted spectral ranges.

Wave optics

The previous considerations only took into account macroscopic systems, i.e. the imaging elements considered were much larger than the wavelength of the observed light beam. Wave optics is concerned with systems in which light is described as an electromagnetic wave. The color of the radiation is defined by the wavelength. Thus, monochromatic light is made up only of light waves of one wavelength, whereas white lights contains light waves of different wavelengths.

Polarization

Polarization is also taken into account when a system is considered from the wave-optics standpoint. This describes the orientation of the waves or the oscillations in relation to the plane of incidence. There are three different types of polarization: linear, circular and elliptical. In the case of linear polarization, the amplitude of the electric field varies with a constant direction of propagation; in the case of circular polarization, the direction varies while the amplitude of the electric field remains constant. Elliptical polarization describes a mixed form of linear and circular polarization.

Linear polarization is further subdivided into s-polarization (perpendicular to the plane of incidence) and p-polarization (in the plane of incidence).

As shown in Figure 6, the characteristic of the Fresnel equations calculated in Figure 2, taking into account s- and

Figure 5: Beam path for a parabolic reflector

A Optical axis,
F Focal point (focus),
P Vertex of parabola.

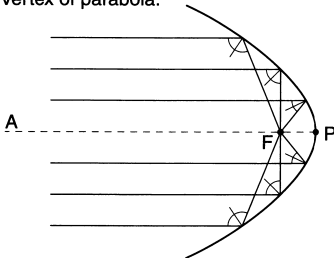
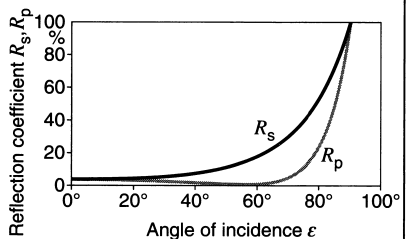


Figure 6: Reflection coefficient

Medium 1 with refractive index $n_1 = 1.00$,
Medium 2 with refractive index $n_2 = 1.50$.



SAN0201E

SAN0200Y

p-polarization, is divided into two different curves. The plane perpendicular to the plane of propagation is called the plane of oscillation. It follows that only waves whose electric field components oscillate in the plane of oscillation are polarizable. These are also called transversal waves (light). In contrast, longitudinal waves (e.g. sound) oscillate in the direction of propagation and are therefore not polarizable.

Interference

The superposition of two waves of equal wavelength can lead to interference. According to the superposition principle, places of constructive and places of destructive interference are formed here. An interference pattern with interference maxima and interference minima is produced. An established example of this is Young's double-slit experiment [4], which proved for the first time the wave nature of light.

Figure 7: Intensity curve $I(x)$ for diffraction at the single slit

L Lens, B Aperture,
D Opening diameter,
x Distance,
 $I(x)$ Intensity profile.

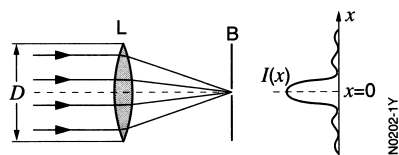
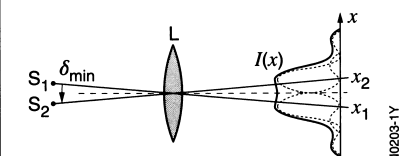


Figure 8: Resolving capacity limited by diffraction of an optical system

L Lens,
 S_1, S_2 Point sources,
 x_1, x_2 Imaging of point sources,
 $\delta_{\min}$ Minimum resolving angle,
 $I(x)$ Intensity profile.



Diffraction

Diffraction phenomena of light waves can also be described using wave optics. The term diffraction is used to refer to the situation where an optical wave encounters an obstacle (e.g. aperture, single slit). The light wave is deflected at the obstacle and new elementary waves are created in accordance with Huygens' principle [4]. These waves interfere with each other and a diffraction pattern is produced in response to interference effects. Thus light penetrates into areas which should be dark according to geometric optics.

Diffraction at a single slit produces in response to interference effects the intensity curve $I(x)$ shown in Figure 7.

When diffraction is considered at an aperture plate or at a circular aperture, the resulting electric field after the aperture can be described by the Bessel functions of the first kind $J_1(2\pi r)$ [5]:

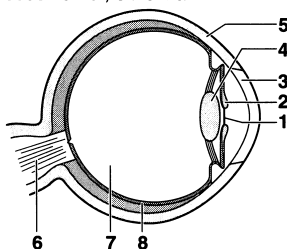
$$E(r) = \frac{E_0}{\pi r} J_1(2\pi r).$$

The resulting intensity distribution after diffraction at the aperture is obtained from

$$I(r) \sim E^2(r).$$

Figure 9: Structure of the eye

1 Pupil, 2 Iris, 3 Cornea, 4 Lens,
5 Sclera, 6 Optic nerve,
7 Vitreous humor, 8 Retina.



Assuming that only very weak (especially the radius) secondary maxima develop as a result of diffraction at a circular aperture, the magnitude of the principal maxima can be determined from the zero point of the Bessel function. This is obtained at $r = 0.6098$.

If diffraction is disregarded, optical systems can achieve an infinite resolution. Diffraction significantly reduces the resolving capacity of any optical system (e.g. lenses, objectives). Resolving capacity U denotes the smallest still perceptible distance between two point sources (S_1 and S_2).

Basically, optical elements are limited by a circular aperture. The resolving capacity of any optical system can thus be calculated from the size of the diffraction pattern of a circular aperture. The intensity curve $I(x)$ is obtained for each point source. The minimum angle $\delta_{\min}$ up to which the points on the screen can still be perceived separately (Figure 8) is calculated as described in the following equation from the wavelength λ of the observed radiation and the opening diameter D :

$$\sin \delta_{\min} = 1.22 \frac{\lambda}{D}.$$

Since the diameter of the aperture opening is calculated from the double radius, the factor 1.22 was determined from the zero point of the Bessel function. The resolving capacity of the system can be determined from $\delta_{\min}$ at $U = 1/\delta_{\min}$.

The human eye can also be considered to be an optical system. In this case the focal plane is the pupil and the shown intensity profile $I(x)$ of a point light source is imaged on the retina (Figure 9). The opening diameter D is determined by the size of the pupil. For a wavelength of $\lambda = 550 \text{ nm}$ and a pupil size of 1.5 mm (daylight vision) $\delta_{\min} = 0.02^\circ$ is obtained for the minimum angle. From this, at a distance of 25 cm to the eye, the minimum resolvable distance between S_1 and S_2 is calculated at 0.11 mm .

Lighting quantities

Light sources have – on the basis of their nature – radiation of different intensities and different wavelengths. The light source is classified by reference to the characteristic wavelength into the wavelength spectrum.

Physical characteristic quantities have become established for the change in intensity of the light source and to describe the effect of light on human beings. A distinction is made here between radiometric and photometric characteristic quantities. In photometry the human eye is evaluated as a detector. This is confined to the ultraviolet (UV) and the visible spectral ranges. In radiometry the power of the light source measured by a detector is determined, thereby also providing measurements in the infrared and ultraviolet ranges and with gamma rays. The lighting quantities in radiometry (radiation quantities) are identified for improved intelligibility by the indices “e” (“e” for energy).

Radiation quantities

Radiant energy Q_e

The radiant energy Q_e comprises the total energy of the light wave.

Radiant flux Φ_e

Radiant flux Φ_e denotes the amount of energy dQ_e that is transported per time unit dt by the light wave:

$$\Phi_e = \frac{dQ_e}{dt}.$$

Irradiance E_e

The proportion of the radiant flux $d\Phi_e$ which strikes a defined surface area dA is called irradiance E_e :

$$E_e = \frac{d\Phi_e}{dA}.$$

The totality of the radiated energy that strikes a bright surface element dA is determined.

Radiant intensity I_e
Radiant intensity I_e denotes the radiant flux $d\Phi_e$ emitted by a point wave in a particular direction (solid angle $d\Omega$).

$$I_e = \frac{d\Phi_e}{d\Omega}.$$

Radiance L_e
Radiance L_e denotes the radiant intensity on a vertical surface.

$$L_e = \frac{dI_e}{dA \cos \varepsilon} = \frac{d^2\Phi_e}{d\Omega dA \cos \varepsilon}.$$

Radiant excitance M_e
Radiant excitance M_e denotes the radiant flux $d\Phi_{e,H}$ of a light source radiated into the half-space:

$$M_e = \frac{d\Phi_{e,H}}{dA}.$$

Irradiation H_e
Irradiation H_e denotes the proportion of the radiant energy dQ that strikes per time dt a surface element dA .

$$H_e = \int E_e dt.$$

Lighting quantities in photometry

Spectral luminous efficiency $V(\lambda)$
Defining the lighting quantities in photometry does not take into account the fact that the eye does not have constant spectral luminous efficiency. The radiation visible to the human eye is in the wavelength range of 380 nm (blue) through

780 nm (red). The eye is at its most sensitive in daylight in the yellow-green range around 555 nm; in weaker light conditions this figure shifts to lower wavelengths. A greater radiant energy is required to achieve the impression of equal brightness for other wavelengths. Spectral luminous efficiency $V(\lambda)$ is defined as the ratio of radiant energy at 555 nm to the radiant energy for the different wavelengths.

Because not all human eyes are identical, a non-dimensional, standardized luminous-efficiency function of the human eye $V(\lambda)$ was defined in DIN 5031-3 [9] for lighting measurements and calculations. This was determined for different wavelength ranges and for daylight and nightlight conditions (Figure 10).

Luminous efficacy of radiation

The absolute spectral efficiency $K(\lambda)$, also called the luminous efficacy of radiation, is the quotient from the physiological quantity luminous flux Φ and the physical radiant flux Φ_e . For daytime vision (eye adapted to the light) the maximum value $K_{m,T}$ of $K(\lambda)$ is obtained at a wavelength of $\lambda = 555$ nm. For nighttime vision (eye adapted to the dark) the maximum $K_{m,N}$ is obtained at $\lambda = 505$ nm. The following relations are obtained for the absolute spectral efficiency $K(\lambda)$:

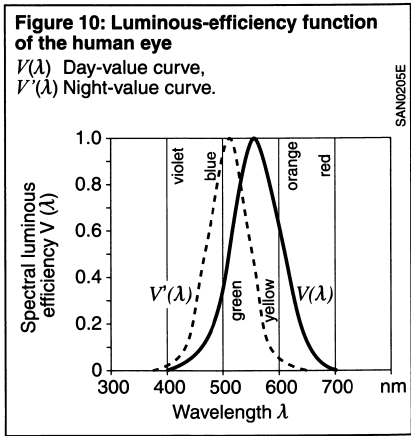


Table 4: Comparison of radiation quantities and photometric quantities			
Radiation quantities		Photometric quantities	
Designation	Unit	Designation	Unit
Radiant energy Q_e	Ws	Quantity of light Q	lms
Radiant flux Φ_e	W	Luminous flux Φ	lm
Radiant intensity I_e	W/sr	Luminous intensity I	cd = lm/sr
Radiance L_e	W/(m ² sr)	Luminance L	cd/m ²
Irradiance E_e	W/m ²	Illuminance E	lx = lm/m ²
Radiant excitance M_e	W/m ²	Luminous excitance M	lx
Irradiation H_e	Ws/m ²	Lumination H	lx s

Daytime vision:

$$K(\lambda) = K_m V(\lambda), \\ K_{m,T} = 683 \text{ lm/W}.$$

Nighttime vision

$$K'(\lambda) = K_m V'(\lambda), \\ K_{m,N} = 1,699 \text{ lm/W}.$$

$K(\lambda)$ enables the radiation quantities X_e and the photometric quantities X to be linked (Table 4). The following applies:

$$X = K_m \int X_{e\lambda} V(\lambda) d\lambda, \\ = \int X_{e\lambda} K(\lambda) d\lambda,$$

$$\text{where } X_{e\lambda} = \frac{dX_e}{d\lambda}.$$

Luminous flux Φ

The luminous flux Φ denotes the weighting of the radiometric radiant flux Φ_e with the wavelength-dependent efficiency curve $K(\lambda)$ of the human eye. The luminous flux Φ determines the light output radiated into the totality of solid angles. All further photometric quantities are linked with the luminous flux.

Quantity of light Q

The quantity of light Q is the photometric equivalent to the radiometric radiant energy Q_e . It is calculated from the integral of the luminous flux Φ over a certain time dt :

$$Q = \int \Phi(t) dt.$$

Luminous intensity I

The luminous flux Φ that is radiated into a certain solid angle Ω is defined as the luminous intensity I . It offers a measure of the luminous emittance of a source radiated in a certain direction and is calculated from the radiometric radiation intensity:

$$I = \frac{\Phi}{\Omega}.$$

Illuminance E

In contrast to luminous intensity I , illuminance E denotes the luminous flux of a light source radiated to a certain surface A :

$$E = \frac{\Phi}{A}.$$

Luminance L

Luminance L defines the brightness sensation in the human eye caused by an illuminated (or also self-luminous) surface. It is obtained from the radiometric quantity for the radiant intensity:

$$L = \frac{I}{\Omega}.$$

Luminous excitance M

The proportion of the luminous flux $d\Phi$ that is generated by a defined surface element dA is known as luminous excitance. It is the photometric counterpart to radiometric radiant excitance:

$$M = \frac{d\Phi}{dA}.$$

Lumination H

The lumination H is calculated from the illuminance E_v that strikes the surface dA over a period of time dt .

$$H = \int E_v(t) dt.$$

Further terms

The lighting quantities are supplemented by the general terms explained in the following.

Luminous efficiency η

Luminous efficiency η is calculated from the ratio of emitted luminous flux Φ to power input P :

$$\eta = \frac{\Phi}{P}.$$

Luminous efficiency is a measure of the efficacy of converting electric power P into luminous flux Φ . It cannot exceed the maximum value of the luminous efficacy of radiation $K_m = 683 \text{ lm/W}$ for the wavelength $\lambda = 555 \text{ nm}$.

Solid angle Ω

The solid angle Ω describes the ratio of the penetrated section of the spherical surface (of a sphere concentric to the radiation source) to the square of the sphere radius. The total surface of the sphere is $4\pi r^2$, thus the full solid angle is

$$\Omega = 4\pi \text{ sr (sr = steradian)}.$$

Steradian is a non-dimensional quantity for the solid angle.

Contrast

The contrast determines the luminance ratio between two neighboring surfaces or the maximum difference in intensity within an illuminated (or self-luminous) surface. The contrast describes the ratio of luminance or the illuminance of a bright area of an image to a dark area of the same image. According to the standard ISO IEC 21118 [10] the contrast must be determined from the values of a black-and-white checkerboard pattern (consisting of sixteen equal squares).

The small-area contrast for alternating black and white lines can also be determined. This is a measure for the specification of an optical lighting system to shown fine details on a screen. Both vertical and horizontal lines are measured for this purpose.

Table 5: Examples of some laser types

Laser type	Wavelength	Applications
He-Ne laser	632 nm	Metrology, holography
CO ₂ laser	10.6 μm	Material processing
ND: YAG laser	1,064 nm	Material processing
Semiconductor laser	e.g. 670 nm e.g. 1,300 nm	Measuring technology Telecommunications
Ytterbium fiber laser	e.g. 1,070 nm	Material processing

Laser technology

Compared to other light sources, the laser (Light Amplification by Stimulated Emission of Radiation) has the following characteristic properties:

- Monochromatic radiation, i.e. a limited wavelength range,
- Very high radiation density,
- Low beam expansion,
- Good focusability,
- High time and spatial coherence of radiation.

Functioning principle

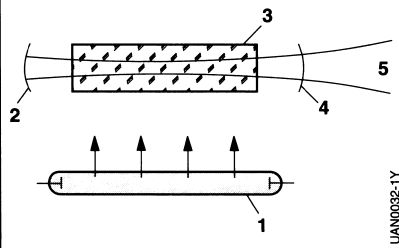
A laser contains a laser-active medium that can be gaseous, solid or liquid (Table 5). The supply of energy can place the atoms or the molecules of the active medium in an excited state (Figure 11). This process is called pumping and can take place electrically (by the application of a voltage) or optically (with another light source).

After a certain period of time the excited particles of the active medium relax back into their initial state and in so doing dissipate their energy through the spontaneous emission of a photon (light particle). The energy of the photon is determined by the quantized energy states of the active medium and stipulates the wavelength λ of the laser light:

$$E_P = \frac{hc}{\lambda}.$$

Figure 11: Laser principle

- 1 Pumping light source,
- 2 Resonator mirror,
- 3 Laser-active material,
- 4 Partly transparent mirror,
- 5 Laser beam.



h is Planck's quantum of action and c the speed of light in a vacuum. For h and c the following values apply:

$$h \approx 6.62606957 \cdot 10^{-34} \text{ Js}$$

$$c \approx 300,000,000 \text{ m/s}$$

The resonator consists of mirror surfaces at the ends of the active medium and causes the spontaneously emitted photons to be reflected back into the active medium. As they pass through the active medium again they cause new photons of equal wavelength and identical phase position to be emitted. This is called stimulated emission.

The resonator is responsible for radiation amplification and for the desired beam characteristic. The laser beam emerges at the end of the resonator via a semi-transparent mirror.

Depending on the active medium and the laser type, lasers can be operated with continuous-wave radiation or in pulsed mode with pulse lengths down to below 1 fs (10^{-15} s) [12, 13].

Areas of application

Laser measuring technology permits the noncontact, non-interacting testing of production tolerances of superfinished surfaces (e.g. fuel injectors). Resolutions in the nm range are achieved using interferometric methods.

Lasers facilitate high-precision, flexible and high-speed material processing/machining in production engineering. For example, hole diameters of 30 μm can be achieved with laser drilling.

Further laser applications in technology are holography (spatial image information), automatic character recognition (bar-code scanners), information recording (CD scanning, 3D space surveying), material processing/machining, microsurgery, and transmitters for data transmission in optical waveguides.

Specific regulations are to be observed when handling laser products. Laser products are classified according to potential hazards. For details, refer to DIN EN ISO 60825 [11].

Optical fibers/waveguides

Design

Optical fibers/waveguides transmit under controlled conditions electromagnetic waves in the ultraviolet (UV), visible, and infrared (IR) ranges of the spectrum. They are made of quartz, glass, or polymers, usually in the form of fibers or channels created in transparent materials with a core whose refractive index is usually higher than that of the cladding. Thus, light launched into the core is retained in that area by means of refraction or total reflection and routed.

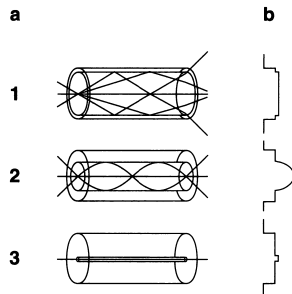
Depending on the refractive index profile, a distinction is made between four types of fiber (Figure 12):

- The step-index fiber, with a sharply-defined boundary between the core and the cladding
- The graded-index fiber, with parabolic refractive index profile in the core
- The monomode fiber with a very small core diameter
- The photonic-crystal fiber with air-filled capillaries arranged periodically around the core. The arrangement corresponds to the step-index or monomode fiber.

Figure 12: Light propagation in optical fibers

a) Fibers schematically represented,
b) Refractive-index profile.

- 1 Step-index fiber,
- 2 Graded-index fiber,
- 3 Monomode fiber.



Step-index and graded-index fibers are multimode fibers, i.e. various oscillation modes of the light waves can be propagated along then at different speeds. The different oscillation modes of the light wave are dependent on the geometry and the material properties of the substrate material. Polymer fibers are always step-index fibers. Monomode fibers are, thanks to their geometry and the material used, designed in such a way that only the fundamental mode of the light wave can be propagated. Depending on the structure configuration, photonic-crystal fibers guide one or more modes.

Properties

Glass optical fibers have a high degree of transparency in the range from ultraviolet to infrared. Losses occur primarily due to the contamination of the fiber material used in the manufacturing process. Thus, H₂O molecules absorbed from the ambient air generate absorption band in the spectrum of the optical fibers. At 950 nm, 1,240 nm and 1,380 nm the transparency of the optical fibers is reduced by the absorption of the H₂O molecules in the glass fiber. When the spectrally resolved absorption of optical fibers is considered, minima are to be seen at 850 nm, 1,310 nm and 1,550 nm, i.e. attenuation is particularly low for the wavelengths 850 nm, 1,310 nm and 1,550 nm. Synthetic fibers absorb above 850 nm and below 450 nm.

Optical fibers can only absorb light from a restricted angular range θ . The numerical aperture $A_N = \sin (\theta/2)$ (Table 6) serves as the measure for this.

The differences in dispersion and propagation time of the various modes cause an increasing broadening of the light pulses as the length of the fiber increases, and thus restrict the bandwidth. In photonic-crystal fibers, it is possible by means of appropriate microstructuring of the core the influence the dispersion and the efficiency of nonlinear effects to achieve desired results.

Optical fibers can be used within the temperature range of -40 to 135°C ; special versions can even be used up to 800°C .

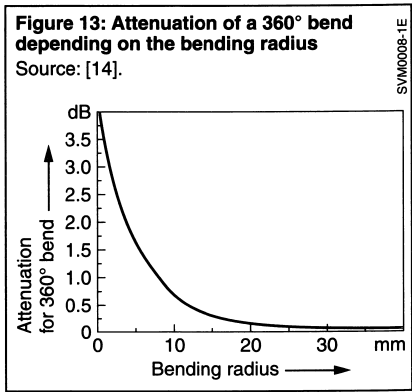


Table 6: Characteristic data of optical fibers/waveguides

Fiber type	Diameter		Wavelength [nm]	Numerical aperture (A_N)	Attenuation [db/km]	Bandwidth [MHz · km]
	Core [μm]	Cladding [μm]				
Step-index fiber	Quartz, glass	50...1,000	70...1,000	250...1,550	0.2...0.87	5...10
	Polymer	200...>1,000	250...2,000	450...850	0.2...0.6	100...500
Graded-index fiber	50...100	100...500	450...1,550	0.2...0.3	3...5	200...10,000
Monomode fiber	3...10	100...500	850...1,550	0.12...0.21	0.3...1	2,500...10,000
Photonic-crystal fiber	1...35	250...200	300...2,000	0.1...0.8	0.2...2	≤ 160,000

Areas of application

The main area of application for optical fibers/waveguides is in data transmission. Synthetic fibers are preferred for use in the LAN (Local Area Network) field. Graded-index fibers are the most suitable for medium ranges. Only mono-mode fibers are used for long-distance data transmission. In fiber-optic networks, erbium-doped glass fibers serve as optical amplifiers. Here, the transported radiation can be boosted by additional optical pumping with a semiconductor source.

Optical fibers/waveguides are used in vehicles in the MOST bus. The required compliance of bending radii means that installation in motor vehicles is critical. If the bending radii are too small, the attenuation is too great (Figure 13, [14]).

Optical fibers are being increasingly used in motor-vehicle lamps and sensors. Fiber-optic sensors generate neither scatter fields nor sparks, and are themselves insensitive to that kind of disturbance. They are currently employed in potentially explosive environments, in medicine, and in high-speed trains (ICE).

Energy transport is at the forefront in the area of material processing with laser beams, in microsurgery, and in lighting engineering.

References

- [1] H. Haferkorn: *Optik – Physikalisch-technische Grundlagen und Anwendungen*. 4th Ed., Wiley-VCH, 2002.
- [2] P.A. Tipler, G. Mosca: *Physik*. 6th Ed.; Springer, 2009.
- [3] W. Demtröder: *Experimentalphysik 2 – Elektrizität und Optik*. 6th Ed.; Springer, 2013.
- [4] G. Litfin: *Technische Optik in der Praxis*. 3rd Ed.; Springer, 2005.
- [5] W. Nolting: *Grundkurs Theoretische Physik 5/2 – Quantenmechanik – Methoden und Anwendungen* (Springer-Lehrbuch), 7th Ed., Springer-Verlag, 2012.
- [6] E. Hecht: *Optik*. 5th Ed., Oldenbourg Wissenschaftsverlag, 2009.
- [7] F. Pedrotti, L. Petrotti, W. Bausch: *Optik – Eine Einführung*. Reihe Prentice Hall, Markt+Technik Verlag, 1996.
- [8] F. Pedrotti, L. Petrotti: *Introduction to Optics*. 3rd Ed., Pearson Education Limited, 2013.
- [9] DIN 5031-3: *Optical radiation physics and illuminating engineering; quantities, symbols and units of illuminating engineering*.
- [10] ISO IEC 21118: *Information technology – Office equipment – Information to be included in specification sheets – Data projectors*.
- [11] DIN EN 60825: *Safety of laser products*.
 Part 1: *Equipment classification and requirements* (2008).
 Part 2: *Safety of optical fibre communication systems* (2011).
 Part 4: *Laser guards* (2011).
 Part 12: *Safety of free space optical communication systems used for transmission of information* (2004).
- [12] W. Radloff: *Laser in Wissenschaft und Technik*. Spektrum Akademischer Verlag, 2011.
- [13] F. K. Kneubühl, M. W. Sigrist: *Laser*. 7th Ed., Vieweg+Teubner, 2008.
- [14] A. Grzempa (Editor): *MOST – Das Multimedia-Bussystem für den Einsatz im Automobil*. 1st Ed., Franzis-Verlag, 2007.

Fluid mechanics

Hydrostatics

Density and pressure

Although fluids are compressible to a lesser extent, they can be viewed as being incompressible for most problems. In addition, since density is only slightly dependent on temperature, it can be taken as constant for many applications.

Pressure $p = dF/dA$ is non-directional in fluids which are at rest. If the pressure component produced by the difference in height (geodetic pressure) is negligible, the hydrostatic pressure is uniformly high everywhere (e.g. in a hydrostatic press).

Fluid at rest in an open vessel

In the case of open vessels, the pressure in the fluid is only dependent on the depth of the fluid (Figure 1). Closed vessels with pressure compensation, such as fuel tanks and brake-fluid reservoirs, can also be considered as open vessels. The following applies:

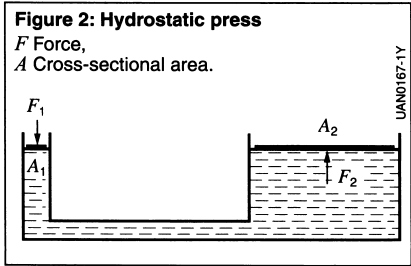
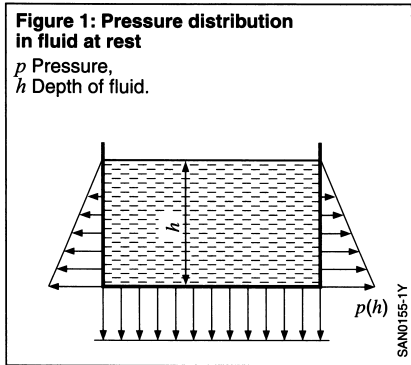
Pressure: $p(h) = \rho gh$
Force acting on bottom: $F_B = A_B \rho gh$
Force acting on sides: $F_S = 0.5 A_S \rho gh$

Hydrostatic press

By way of example, power amplification in vehicle brakes and hydraulic power-assisted steering systems functions according to the hydrostatic-pressure principle (Figure 2). The following relationships apply to the pressure p and the piston forces F :

Table 1: Symbols and units

Quantity	Unit
A Cross-sectional area	m^2
A_B Area of base	m^2
A_S Area of side	m^2
F Force	N
F_A Buoyancy force	N
F_B Force acting on bottom	N
F_G Weight	N
F_S Lateral force	N
F_W Volume of displaced fluid	N
L Length in flow direction	m
Q Volumetric flow	m^3/s
Re Reynolds number	–
V_F Volume of displaced fluid	m^3
c_W Drag coefficient	–
d Diameter	m
g Acceleration of free fall ($g \approx 9,81 \text{ m/s}^2$)	m/s^2
h Depth of fluid	m
m Mass	kg
m_F Mass of displaced fluid	kg
$\dot{m}$ Mass flow	kg/s
p Pressure	$Pa = N/m^2$
t Thickness	m
w Flow velocity	m/s
α Contraction coefficient	–
η Dynamic viscosity	$Pa \cdot s = Ns/m^2$
μ Discharge coefficient	–
ν Kinematic viscosity	m^2/s
ρ Density	kg/m^3
φ Velocity coefficient	–
τ Shear stress	N/m^2



Pressure:
$$p = \frac{F_1}{A_1} = \frac{F_2}{A_2}.$$

Piston forces:
$$F_1 = pA_1 = F_2 \frac{A_1}{A_2},$$

$$F_2 = pA_2 = F_1 \frac{A_2}{A_1}.$$

Buoyancy

Buoyancy is a force acting against gravity and acts on the center of gravity of the volume of the displaced fluid. It corresponds to the weight of the fluid displaced by the submerged body:

$$F_A = m_F g = V_F \rho g.$$

A body will float if $F_A = F_G$.

The amount of fuel available can be easily and reliably measured by analog sensors (floats) in the fuel tank with the aid of buoyancy.

Flow mechanics

Basic principles

An ideal fluid (generic term for gases and liquids) is incompressible and frictionless. This means that no shear stresses occur in the fluid, and the pressure on a fluid element is uniform in all directions. In actual fact, however, a resistance must be overcome in fluids if deformations occur that are caused by displacement of fluid elements (Figure 3). The resulting shear stress complies with Newton's:

$$\tau = \frac{F}{A} = \eta \frac{w}{h}.$$

The proportionality factor η is called the dynamic viscosity and is greatly dependent on temperature. In practice, however, the kinematic viscosity

$$\nu = \frac{\eta}{\rho}$$

is often used, as it can be measured very easily with a capillary viscometer.

Flows without turbulence, in which the individual fluid layers move separately in parallel and which are predominantly determined by viscosity, are known as laminar flows. If the flow velocity exceeds a limit value, adjacent layers start to swirl, resulting in a turbulent flow. In addition to the flow velocity, the transition point between laminar and turbulent flow is also dependent on the Reynolds number

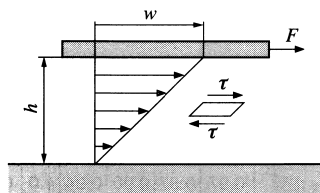
$$Re = \frac{\rho L w}{\eta} = \frac{L w}{\nu}$$

In the case of a flow within a pipe, the pipe diameter is used for L . Flow in a pipe becomes unstable or turbulent at $Re > 2300$.

Since, in the case of gases with low flow velocities (up to 0.5 times the velocity of sound), compression is negligible in many flow processes, they are also governed by the laws of incompressible fluids.

Figure 3: Shear stresses in fluids

τ Shear stress,
 w Flow velocity,
 h Height,
 F Force.



SAND157-2Y

Basic equations of flow mechanics

The most important basic equations of flow mechanics are the continuity equation and the Bernoulli equation. They describe the conservation of mass and energy in flowing fluids.

Continuity equation

In a steady state, mass conservation requires that in a flow the mass flow rate be of equal magnitude in each cross-section (Figure 4):

$\dot{m} = \rho A_1 w_1 = \rho A_2 w_2 = \text{const.}$

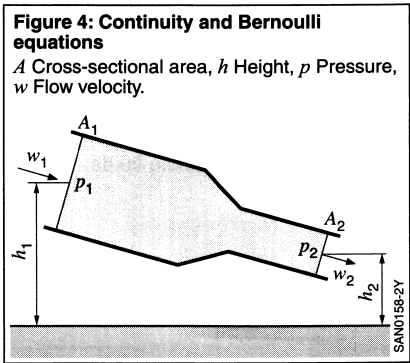
In the case of incompressible fluids ($\rho = \text{const.}$), the volumetric flow must also be constant:

$Q = A_1 w_1 = A_2 w_2 = \text{const.}$

Bernoulli equation

From the continuity equation, it follows that an acceleration takes place between A_1 and A_2 . This results in an increase in kinetic energy, which must be effected by a pressure drop, where $p_1 > p_2$ (Figure 4). According to the law of conservation of energy, the sum of the static pressure p , kinetic pressure, and geodetic pressure is constant in a flowing fluid. Ignoring friction losses, the following applies accordingly to the flowing fluid in a non-horizontal pipe:

$p_1 + \frac{1}{2} \rho w_1^2 + \rho g h_1 = p_2 + \frac{1}{2} \rho w_2^2 + \rho g h_2 .$



Discharge from a pressure vessel

Under the precondition that the cross-sectional area of the discharge end is very much smaller than that of the vessel (Figure 5), the velocity w_1 is negligible according to the continuity equation. As derived from Bernoulli equation, the discharge velocity is governed by the following:

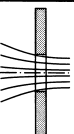
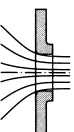
$w_2 = \varphi \sqrt{\frac{2}{\rho} (p_1 - p_2) + 2 g h} .$

The velocity coefficient φ takes into account the losses that occur. The jet contraction must also be taken into account for the volumetric flow or the discharge volume; this constriction is dependent on the contraction coefficient α . The following then applies to the discharge volumetric flow:

$Q = \alpha \varphi A_2 \sqrt{\frac{2}{\rho} (p_1 - p_2) + 2 g h} .$

Velocity coefficient and contraction coefficient are often expressed together as the discharge coefficient $\mu = \alpha \varphi$ (Table 2).

Table 2: Discharge openings

Orifice shape	Velocity coefficient φ	Contraction coefficient α	Discharge coefficient μ									
	0.97	0.61...0.64	0.59 ... 0.62									
	0.97 ... 0.99	1.0	0.97 ... 0.99									
	0.95 ... 0.97	$(d_2/d_1)^2$ <table><tr><th>0.4</th><th>0.6</th><th>0.8</th><th>1.0</th></tr><tr><td>0.87</td><td>0.90</td><td>0.94</td><td>1.0</td></tr></table>		0.4	0.6	0.8	1.0	0.87	0.90	0.94	1.0	0.82 ... 0.97
0.4	0.6	0.8	1.0									
0.87	0.90	0.94	1.0									

Resistance of bodies submerged in a fluid flow

A pressure differential occurs across a body submerged in a fluid flow (e.g. the vehicle body), resulting in a resistance force

$$F_W = \frac{1}{2} c_W A \rho w^2.$$

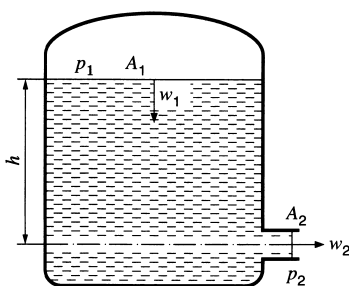
Here, A is the cross-sectional area of the body on which the fluid flow is acting and c_d an undefined coefficient of resistance, which is dependent on the shape of the body submerged in the fluid flow.

As it is extremely complex to calculate exactly the resistance to flow even for simple bodies, resistance to flow is usually determined experimentally. If the dimensions are large, the measurements are taken on downscaled models. As well as geometrical similarity, the forms of energy that occur (kinetic energy, frictional work) in the original fluid flow and in the model flow must be proportional. This proportion is denoted by the Reynolds number R_e .

Basically: Two flows are similar in fluid-dynamic terms if their Reynolds numbers R_e are identical. Because even complex geometries are composed of simple basic bodies, streamlined surfaces can already be derived during the modeling phase as per Table 3.

Figure 5: Discharge from a pressure vessel

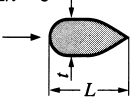
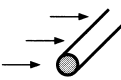
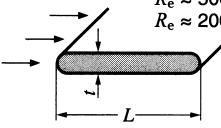
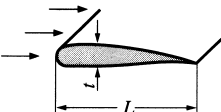
A Cross-sectional area, h Height, p Pressure, w Flow velocity.



References

- [1] A. Böge; W. Böge: Technische Mechanik. 31st Ed., Verlag Springer Vieweg, 2015.
[2] W. Bohl; W. Elmendorf: Technische Strömungslehre. 15th Ed., Vogel-Verlag, 2014.

Table 3: Drag coefficients c_W

Body shape: L Length, t Thickness, R_e Reynolds number.	c_W
Circular plate 	1.11
Open dish 	1.33
Sphere  $R_e < 200,000$ $R_e > 250,000$	0.47 0.20
Narrow rotational body $L/t = 6$ 	0.05
Long cylinder  $R_e < 200,000$ $R_e > 450,000$	1.0 0.35
Long plate $L/t = 30$  $R_e \approx 500,000$ $R_e \approx 200,000$	0.78 0.66
Long wing $L/t = 18$ $R_e \approx 10^6$ $L/t = 8$ $R_e \approx 10^6$ $L/t = 5$ $R_e \approx 10^6$ $L/t = 2$ $R_e \approx 2 \cdot 10^5$ 	0.2 0.1 0.08 0.2

SAN0161-2Y

Thermodynamics

Basic principles

System

Thermodynamic systems are typically divided into three types. What all the systems have in common is that they are delimited from the surroundings by the system boundary within which the system is located.

Isolated system

The system boundary of an isolated system is impermeable to all types of heat Q , work W and mass m (e.g. the medium inside a perfectly insulated Thermos flask).

Closed system

A closed system allows the exchange of heat and work with the surroundings across the system boundary, but not of mass (e.g. a gas in a cylinder sealed by a moving piston).

Open system

Finally, in an open system heat, work and mass can be exchanged across the system boundary with the surroundings (e.g. an open cooking pot).

State and process variables

State variables

In thermodynamics states of systems and (changes to) processes are described by way of state variables such as for example temperature T , pressure p , volume V or mass m . The changes to these variables are called changes of state. The state of a system is described uniquely by means of the state variables. To describe the state of a system which undergoes a change of state from state 1 to state 2, it is therefore not necessary to know the path taken between initial and final states, but solely to know the relevant state variables.

Process and process variables

Generally speaking, the stringing together of changes of state is called a process. In contrast to state variables, the process variables required to describe the process depend on the path between initial and final states. Work and heat are thus process variables, the type and with it also the magnitude of the work input differ e.g. depending on process control.

Table 1: Symbols and units

Quantity	SI unit
A Cross-sectional area	m^2
a Thermal diffusivity	m^2/s
c Specific heat capacity	$\text{J}/(\text{kg} \cdot \text{K})$
c_p Isobaric (constant pressure)	
c_v Isochoric (constant volume)	
E Energy	J
e Specific energy	J/kg
H Enthalpy	J
h Specific enthalpy	J/kg
k Heat-transmission coefficient	$\text{W}/(\text{m}^2 \cdot \text{K})$
m Mass	kg
n Polytropic exponent	—
p Pressure	$\text{Pa} = \text{N}/\text{m}^2$
Q Heat	J
$\dot{Q}$ Heat flow dQ/dt	W
R_m Universal gas constant	$\text{J}/(\text{mol} \cdot \text{K})$
$R_m \approx 8.3145 \text{ J}/(\text{mol} \cdot \text{K})$	
R_i Specific gas constant	$\text{J}/(\text{kg} \cdot \text{K})$
$R_i = R_m/M$ (M molar mass)	
R_λ Thermal-conduction resistance	K/W
S Entropy	J/K
s Length	m
T Thermodynamic temperature	K
ΔT Temperature difference	K
$\Delta T = T_1 - T_2$	
U Internal energy	J
u Specific internal energy	J/kg
V Volume	m^3
v Specific volume	m^3/kg
W Work	J
W_t Technical work	J
t Time	s
α Heat-transfer coefficient	$\text{W}/(\text{m}^2 \cdot \text{K})$
ε Emissivity	—
κ Isentropic exponent	—
λ Thermal conductivity	$\text{W}/(\text{m} \cdot \text{K})$
ρ Density	kg/m^3
ν Kinematic viscosity	m^2/s
σ Stefan-Boltzmann constant	$\text{W}/\text{m}^2\text{K}^4$
$\sigma \approx 5.6704 \cdot 10^{-8} \text{ W}/\text{m}^2\text{K}^4$	

When a system runs through a sequence of processes and ends again in the initial state, this is called a cycle. Cycles are important for describing many technical applications (e.g. engines, power plants, power stations, air conditioners).

Intensive and extensive state variables

State variables can be further subdivided into intensive and extensive variables. When an existing system is subdivided into two parts, all the variables whose values are maintained in the two new systems are called intensive variables (e.g. temperature, pressure). Variables whose values change are called extensive variables (e.g. volume).

Often of interest are system properties which do not depend on the absolute value of a system. It can therefore be practicable to divide the extensive variables by the system mass. This results in specific state variables which retain their value when the system is subdivided (e.g. specific volume $v = V/m$ and its reciprocal value, density $\rho = 1/v$).

Forms of energy

As in classic mechanics, the state variables “kinetic energy” and “potential energy” are defined in thermodynamics. When mass m moves at constant speed c , it thus has the kinetic energy

$$E_{\text{kin}} = \frac{1}{2} m c^2.$$

A mass m that is under gravity at the acceleration of free fall g and height z has the potential energy

$$E_{\text{pot}} = m g z.$$

In addition to these forms of energy, a thermodynamic system has a further inherent type of energy, internal energy U . As well as the movement of the system, this type of energy furthermore takes into account changes within the system due for example to a change in temperature. The total energy of a system is then:

$$E_{\text{tot}} = U + E_{\text{kin}} + E_{\text{pot}}.$$

Work and heat

Work W is generally defined as the integral of the force F acting at a point of application on the system via the distance covered ds (from s_1 to s_2):

$$W_{12} = \int_{s_1}^{s_2} F ds.$$

The special case of volume-change work is crucially important in thermodynamics. This is defined as

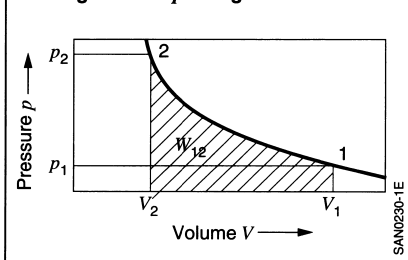
$$W_{12} = - \int_{V_1}^{V_2} p dV.$$

Volume-change work can be shown and read as the area (integral) under the curve in a pV diagram (Figure 1).

In contrast to work, heat Q is a form of energy that occurs in an unordered way. When energy is supplied to a system for example by an electric heater, the change in the system's state of equilibrium is caused by heat, but not by work, which in this example is equal to zero.

The two variables work and heat are counted positively when they are supplied to a system. When work or heat is rejected, they are negative. This equates to a system-egotistical standpoint.

Figure 1: Volume-change work as integral in the pV diagram



Laws of thermodynamics

Zeroth law

Two systems which are each in thermal equilibrium with a third system are also in thermal equilibrium with each other. All the systems thus have the same temperature.

First law

The first law of thermodynamics states that energy cannot be created or destroyed, but only converted from form to another, e.g. heat to work.

For an isolated system, in which there are no flows of mass or energy across the system boundaries, the sum of all the energy changes equals zero.

For a closed system, in which energy but not mass can flow across the system boundaries, the following applies:

$$Q_{12} + W_{12} = \Delta U + \Delta E_{\text{kin}} + \Delta E_{\text{pot}}.$$

The sum of the change in the internal energy ΔU and the kinetic energy ΔE_{kin} , and the potential energy ΔE_{pot} of the system is, therefore, the sum of the quantity of heat Q_{12} added and the work W_{12} done on the system (in the case of a change from state 1 to state 2).

In many technical applications a medium flows through an open system in a stationary flow process. It is therefore also necessary to consider that the intake and expulsion of the input or output mass is connected with work W_1 and the input or output mass has an additional energy content. For a stationary flow process with an input and an output mass flow

$\dot{m}_1 = -\dot{m}_2 = \dot{m}$ the following applies:

$$\dot{Q}_{12} + \dot{W}_{t,12} = \dot{m} (\Delta h + \Delta e_{\text{kin}} + \Delta e_{\text{pot}}),$$

with the new state variable enthalpy:

$$h = u + p v.$$

e denotes the specific variable, i.e.:

$$e = \frac{E}{m}.$$

A perpetual mobile of the first kind violates the first law of thermodynamics in that it creates “energy from nothing” or presupposes an (impossible) efficiency of more than 100 %.

Second law

Once the retention of the total energy has been described by the first law, the second law of thermodynamics specifies the direction of the energy exchange and the sequences of processes. For example, heat transfer can only take place from a hot body to a cold body, never the other way round. The decisive state variable is entropy. For reversible changes of state the change dS in the entropy S of a system is defined as:

$$dS = \frac{dQ}{T}.$$

In other words, the transferred quantity of heat dQ is divided by the temperature T at the point of heat exchange. If therefore a quantity of heat dQ is isothermally added for example to a system, this results in an increase in the statistical fluctuation of the kinetic energy of the atoms (effectively the addition of heat) and also an increase in its Entropy by the value dS .

For all irreversible changes of state and thus for all real technical processes (accompanied by loss) it follows as a result of energy dissipation that entropy production is positive. In the case of a reversible process, however, no entropy is created. To summarize, it can therefore be stated that the rate of entropy production is never negative and thus:

$$dS \geq 0.$$

A perpetuum mobile of the second type violates the second law of thermodynamics, while it can perfectly comply with the first law. The second main law states that it is not possible to convert heat into work when the ambient temperature stays the same. Nor is it possible to convert a quantity of heat which exists at a temperature level above the surroundings fully into work.

Third law

The third law assigns to the entropy at absolute zero, at which the temperature $T = 0$ K, a value that is not dependent on pressure, temperature, volume, etc. This is $S_0 = 0$ J/K.

It also states that absolute zero can be approached by a series of different processes only asymptotically, but never reached.

Table 2: Specific heat capacity c_p at constant pressure for some gases

Values apply to 1 bar and 293.15 K [1].

Gas	c_p [kJ/kg·K]
Nitrogen (N ₂)	1.041
Oxygen (O ₂)	0.9189
Helium	5.251
Air	1.007
Carbon dioxide	0.8459
Ammonia (NH ₃)	2.160

Equations of state

The thermal behavior of a gas is described by way of the correlation of pressure p , temperature T and volume V , i.e. by way of a function $F(p, V, T) = 0$. This function is called the thermal equation of state.

As well as describing the thermal behavior, a further equation, which establishes a correlation between the internal energy U and the thermal state variables, is needed to describe the caloric behavior. The caloric equation of state thus reads $U = U(V, T)$. This correlation is required when for example from the first law the resulting temperature change is to be calculated after a certain quantity of heat has been added.

Ideal gas

The thermal equation of state for an ideal gas reads:

$$pV = mRT \quad \text{or specifically} \\ pv = RT,$$

with the (specific) gas constant R ($R = R_m/M$) and the mass m . In addition, the following applies:

$$u = u(T) \text{ and } h = h(T).$$

In other words, the internal energy u and the enthalpy h of an ideal gas only depend on the temperature T . From the differentials of the internal energy and the enthalpy it follows with $(du/dT)_v = c_v$ and $(dh/dT)_p = c_p$:

$$u(T) = \int_{T_0}^T c_v(T) dt + u_0,$$

$$h(T) = \int_{T_0}^T c_p(T) dt + h_0.$$

This correlation can be further simplified when c_v and c_p are assumed to be constant, i.e. independent of the temperature. Such a gas is called a perfect gas. Ideal gases are governed by the correlation

$$c_p - c_v = R.$$

Table 2 shows by way of example values for c_p of some gases.

Real gas

The ideal gas equation demonstrates in certain ranges, e.g. at very high pressures, significant deviations from reality. An equation which better describes this behavior is the van der Waals equation:

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT.$$

Here a and b are substance-specific variables, taking into account the influence of internal pressure a/v^2 , resulting from molecular attraction, and the specific volume b of the molecules. There are a large number of other real gas equations as well as this equation. For these and further details, refer to the further literature [2], [3], [4].

Table 3: The most important changes of state of ideal gases and their properties

Change of state	Diagram	Equations for W_{12} volume-change work and Q_{12} added heat. (For quantities used, see Table 1).
Isochoric ($V = \text{const}$)		$W_{12} = 0$ $Q_{12} = m c_v (T_2 - T_1)$
Isobaric ($p = \text{const}$)		$W_{12} = -p (V_2 - V_1)$ $Q_{12} = m c_p (T_2 - T_1)$
Isothermal ($T = \text{const}$)		$W_{12} = -p_1 V_1 \ln \frac{p_1}{p_2}$ $Q_{12} = -W_{12}$
Reversible adiabatic ($dS = 0$)		$W_{12} = \frac{p_1 V_1}{\kappa - 1} \left(\left(\frac{V_1}{V_2} \right)^{\kappa - 1} - 1 \right)$ $Q_{12} = 0$
Polytropic curve ($p v^n = \text{const}$)		$W_{12} = \frac{p_1 V_1}{n - 1} \left(\left(\frac{V_1}{V_2} \right)^{n - 1} - 1 \right)$ $Q_{12} = m c_v \frac{n - \kappa}{n - 1} (T_2 - T_1)$

Changes of state

Real changes of state of gases can be approached for many technical processes by simplifying, idealized assumptions. The technically most important changes of state include the isothermic ($T = \text{const}$), the isobaric ($p = \text{const}$), the isochoric ($V = \text{const}$), and the reversible adiabatic ($\dot{q} = 0$ or $dS = 0$, no heat exchange). In addition to description by these simplified idealized changes of state, which are often realized to a good degree of approximation, it is possible, assuming an ideal gas, to describe many technical processes with greater accuracy by means of a polytropic change of state of the form

$$p V^n = \text{const}.$$

Depending on the choice of exponent n , different process paths are reproduced, thus the mentioned isothermic ($n = 1$), isobaric ($n = 0$), isochoric ($n \rightarrow \infty$), and reversible adiabatic ($n = \kappa$, with the isentropic exponent κ) changes of state. The changes of state, the associated correlations of the variables p , V and T , the associated equations for volume-change work

$$W_{12} = - \int_{V_1}^{V_2} p dV$$

and the added or rejected heat Q_{12} are set out in Table 3. Work and heat can further be read from the illustrated pV and TS diagrams as areas.

Cycles and technical applications

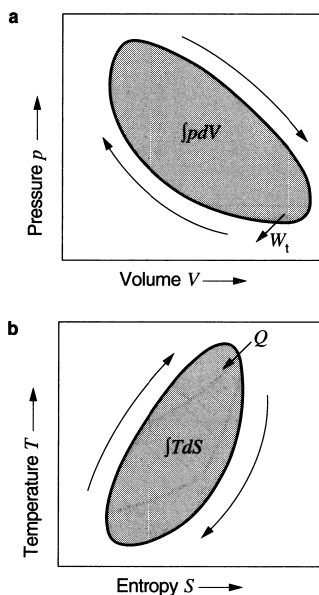
Basic principles

A cycle is a sequence of different thermodynamic changes of state (processes) which when completed sees the return of the initial state, i.e. the state variables in the initial and final states are identical. Here, a cycle can take place with a constant, circulating mass flow in a closed system (e.g. in a Stirling engine or a refrigerating machine) or in an open system with media exchange (e.g. in an internal-combustion engine or a gas turbine).

A cycle is typically such that either it is deprived of work (as in a heat engine) or the heat is increased by the addition of work to a higher temperature level (as in a refrigerating machine or a heat pump). These two forms can be subdivided with regard to the successive direction of their changes of state in the pV and TS diagrams into clockwise (engines) and counterclockwise machines (machines). The

Figure 2: Clockwise cycle

- a) Depiction in the pV diagram,
b) Depiction in the TS diagram.



areas enclosed by the cycle represent here the input work in the pV diagram and the added heat in the TS diagram (Figure 2).

Therefore heat is added and work is output net for the situation of a clockwise cycle shown in Figure 2. The efficiency η of such a process is usually represented as the ratio of benefit (work output W_i) and expenditure (heat addition Q_{ad}), in this case therefore:

$$\eta = \frac{|W_i|}{|Q_{ad}|}.$$

For technical applications the real sequences are approached by comparison cycles. Here the entire cycle is subdivided into different reversibly accepted subcycles. The latter can be described by means of simple changes of state (e.g. isentropic, isothermal, isobaric, isochoric). In addition to simple descriptibility, such idealized comparison cycles can be used as reference for comparing real machines as they represent the optimum in the respective process control.

Carnot cycle

The Carnot cycle constitutes the ideal comparison cycle which shows the best possible efficiency of a machine operating between two temperature levels ($T_{\min}$ and $T_{\max}$). Heat addition and heat rejection occur in the Carnot cycle isothermally as shown in Figure 3 and thus without dissipation, work is done (compression and expansion) in reversible adiabatic fashion.

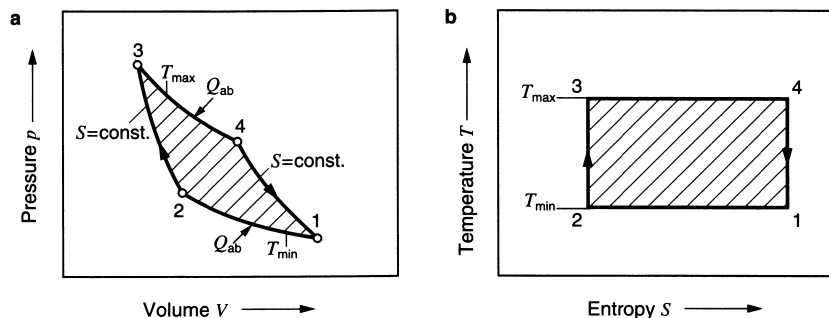
For Carnot efficiency the following equation is obtained for an ideal gas according to the above definition and by applying the first law:

$$\eta = \frac{|W_i|}{|Q_{ad}|} = 1 - \frac{T_{\min}}{T_{\max}}.$$

The Carnot cycle has extra significance because – although in reality it can only be approximately represented at great expense – it describes the describes the maximum amount of utilizable energy (exergy) or non-utilizable energy (anergy) and the maximum possible efficiency can be achieved with it.

Figure 3: Changes of state of the Carnot cycle

a) In the pV diagram, b) In the TS diagram.



Comparison cycles of technical applications

Seiliger cycle

The Seiliger cycle describes the sequences and processes in internal-combustion engines. It also depicts the Diesel cycle (constant-pressure cycle) and the Otto cycle (constant-volume cycle) as boundary cycles. In the Seiliger cycle heat rejection takes place through combustion. This can be taken as isochoric (with a virtually stationary piston and constant volume) or isobaric (with a moving piston and constant in-cylinder pressure). The two forms and mixed forms of these two forms are depicted in the Seiliger cycle. The following five cycles steps are featured and depicted in Figure 4 in the pV and in the TS diagrams:

- Reversible adiabatic compression (1→2),
- Isochoric heat addition (2→3),
- Isobaric heat addition (3→4),
- Reversible adiabatic expansion (4→5),
- Isochoric heat rejection (5→1).

Important characteristics of the Seiliger cycle are the compression ratio $\varepsilon = V_1/V_2$, the pressure-increase ratio $\psi = p_1/p_2$, and the injection ratio $\phi = V_4/V_3$. For the thermal efficiency of the Seiliger cycle the following expression is obtained for an ideal gas:

$$\eta_{th} = 1 - \varepsilon^{1-\kappa} \frac{\psi \phi^{\kappa} - 1}{\psi - 1 + \kappa \psi(\phi - 1)}.$$

Figure 5 shows the efficiencies that can be achieved of different process controls as a function of compression and pressure-increase ratios. It also shows the values of the boundary cases of constant-volume and constant-pressure cycles. These two

Figure 5: Efficiencies of the comparison cycles as a function of compression ratio

- 1 Constant-volume cycle,
 - 2 Seiliger limit-pressure cycle, $\phi = 1.5$, $\psi = 5.0$,
 - 3 Seiliger limit-pressure cycle, $\phi = 1.5$, $\psi = 1.5$,
 - 4 Constant-pressure cycle, $\phi = 1.5$,
 - 5 Seiliger limit-pressure cycle, $\phi = 2.0$, $\psi = 5.0$,
 - 6 Seiliger limit-pressure cycle, $\phi = 2.0$, $\psi = 1.5$,
 - 7 Constant-pressure cycle, $\phi = 2.0$,
- ε Compression ratio,
 ϕ Injection ratio,
 ψ Pressure-increase ratio.

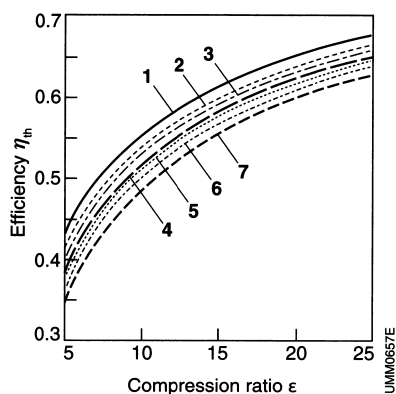
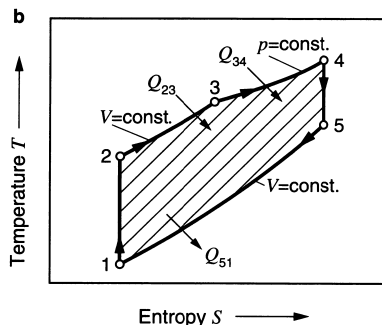
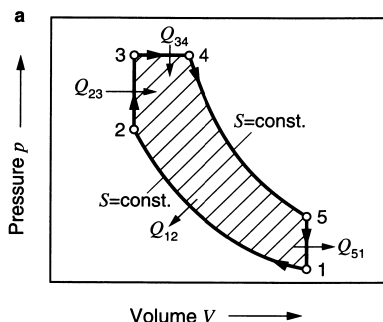


Figure 4: Changes of state of the Seiliger cycle

a) In the pV diagram, b) In the TS diagram.



boundary cases are described in further detail in the following.

Otto cycle

The Otto cycle (constant-volume cycle) represents the boundary process of the Seiliger cycle, in which the addition of heat or combustion takes place exclusively isochorically. Compared with the Seiliger cycle, therefore, the cycle step of isobaric heat addition ($3 \rightarrow 4$) is absent. Process control in the pV and TS diagrams is illustrated in Figure 6. The following simplified description is obtained for the efficiency of the constant-volume cycle:

$$\eta_{\text{th}} = 1 - \varepsilon^{1-\kappa}.$$

The efficiency increases as the compression ratio ε increases. In gasoline engines the compression ratio is limited by knocking. Naturally aspirated gasoline engines operate at compression ratios of $\varepsilon = 10$ – 12 , the maximum pressures are approximately 60 bar. Turbocharged engines, owing to the knock limit, operate at lower compression ratios, e.g. at $\varepsilon = 9$ – 10 . Peak pressures of up to 120 bar are reached. Direct-injection stratified-charge engines have compression ratios $\varepsilon > 12$; at part load the potential for variable compression ratios is even up to $\varepsilon = 14$.

Diesel cycle

The second boundary process of the Seiliger cycle is the Diesel cycle (constant-pressure cycle). Process control in the pV and TS diagrams is illustrated in Figure 7. In this cycle the addition of heat or combustion takes place isobarically, i.e. the cycle step of isochoric heat addition ($2 \rightarrow 3$) is absent. The efficiency of the constant-pressure cycle is obtained as:

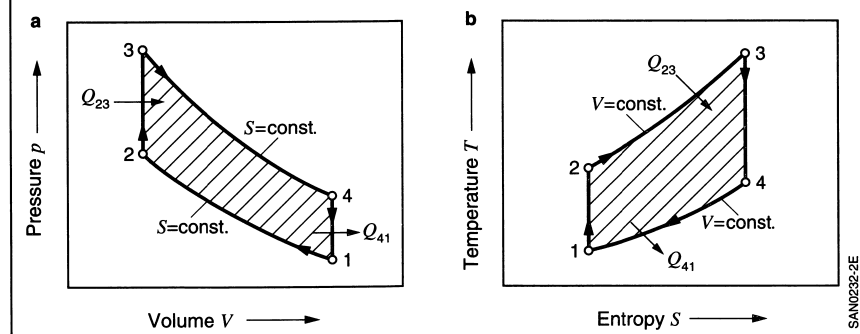
$$\eta_{\text{th}} = 1 - \frac{\varphi^{\kappa} - 1}{\varepsilon^{\kappa-1} \kappa (\varphi - 1)}.$$

The achieved efficiency is, for the same compression ratio, lower than in the constant-volume cycle ε . However, because diesel engines are typically operated with a higher compression ratio, their efficiency is generally better. The aim of development for diesel engines is therefore to deliver a high peak pressure.

The maximum permissible peak pressures for passenger-car engines are roughly 180 bar, and for commercial-vehicle engines over 220 bar. The compression ratios for the direct-injection processes currently used today are $\varepsilon = 16$ – 19 . In the case of retarded start of injection, the final compression pressure actually corresponds to the peak pressure, and the diesel cycle resembles the constant-pressure cycle with $\varphi \approx 9$.

Figure 6: Changes of state of the Otto cycle

a) In the pV diagram, b) In the TS diagram.



Joule cycle

The Joule cycle (or Brayton cycle) is the comparison cycle for gas turbines. In contrast to internal-combustion engines and their comparison cycles, a gas turbine is subject to continuous flow. The Joule cycle is characterized by the following cycle steps (Figure 8):

- Reversible adiabatic compression ($1 \rightarrow 2$),
- Isobaric heat addition ($2 \rightarrow 3$),
- Reversible adiabatic expansion ($3 \rightarrow 4$),
- Isobaric heat rejection ($4 \rightarrow 1$).

An important characteristic of the Joule cycles is the pressure ratio $\pi = p_1/p_2$. The efficiency is thus:

$$\eta_{\text{th}} = 1 - \frac{T_1}{T_2} = 1 - \pi^{\frac{\kappa-1}{\kappa}}.$$

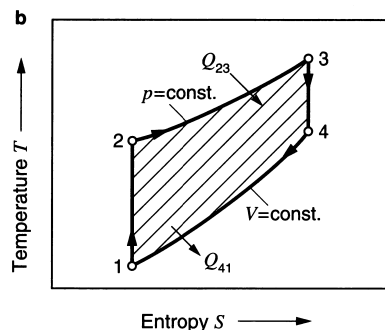
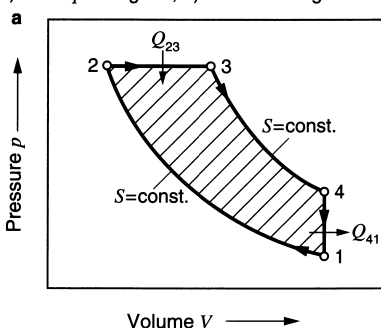
Clausius-Rankine cycle

The Clausius-Rankine cycle is usually used as the comparison cycle for steam turbines. This is a closed cycle in which the working medium undergoes among others two phase changes (evaporation and condensation). The cycle typically contains the following steps (Figure 9):

- Reversible adiabatic pressure increase in the liquid phase ($1 \rightarrow 2$),
- Isobaric heat addition ($2 \rightarrow 3$),
- Isobaric and isothermal evaporation ($3 \rightarrow 4$),
- Isobaric overheating ($4 \rightarrow 5$),
- Reversible adiabatic expansion ($5 \rightarrow 6$),
- Isobaric and isothermal condensation ($6 \rightarrow 1$).

Figure 7: Changes of state of the Diesel cycle

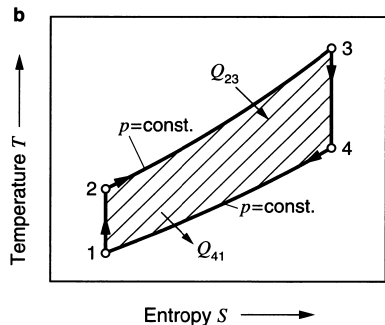
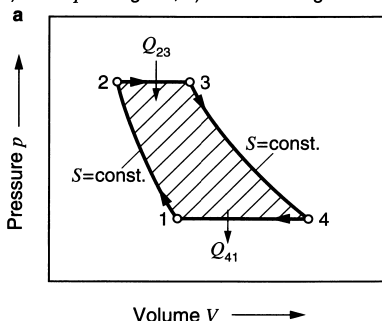
a) In the pV diagram, b) In the TS diagram.



SAN0234-2E

Figure 8: Changes of state of the Joule cycle

a) In the pV diagram, b) In the TS diagram.



SAN0234-2E

The two curves in Figure 9 relate to the phase boundary lines (saturated-liquid and saturated-vapor lines). The saturated-liquid line represents the range of the just boiling liquid, the saturated-vapor line that of the saturated vapor. Below these two lines is the wet-steam region, i.e., the region in which liquid and gaseous phases exist together. The summit of the curves is called the critical point.

The Clausius-Rankine cycles on account of the phase changes cannot be depicted with an ideal gas; it requires a real-gas description. For the same reason, this produces a depiction which describes benefit and expenditure and thus the thermal efficiency with enthalpies. The thermal efficiency is thus [3]:

$$\eta_{\text{th}} = 1 - \frac{h_6 - h_1}{h_5 - h_2}.$$

In automotive applications the Clausius-Rankine cycle is used for example in the field of waste-heat recovery as a comparison cycle. The engine or exhaust-gas waste heat supplies the required heat here for the cycle steps 2 → 5. A turbine or a piston engine for example which makes energy available in mechanical form can be used as the engine for expansion (5 → 6).

Heat transfer

Concept formation

Essentially, heat is transferred in three different ways.

Thermal conduction

In the case of thermal conduction, heat is transferred on the molecular level. Heat is transferred here as a result of an impressed temperature gradient by a solid body or a static medium (liquid or gaseous) without macroscopic matter transport. According to the second law of thermodynamics heat is transferred here in the direction of the lower temperature.

Convective heat transfer

In the case of convective heat transfer, a flow of matter, as exists only in fluid, flowing media (liquids or gases), is required. A further distinction is made between free (caused by natural buoyancy on account of the prevailing temperature gradients) and forced (caused by externally impressed flows) convection.

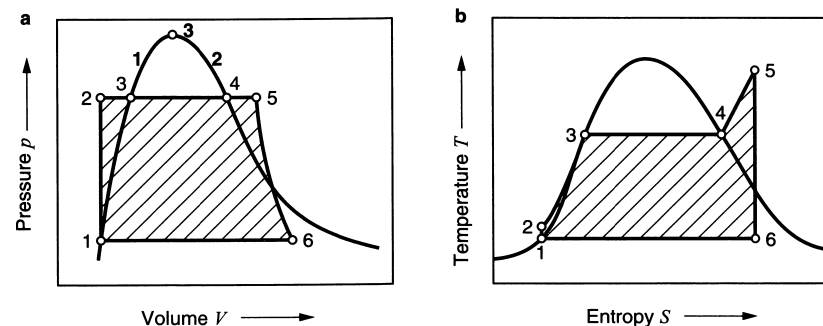
Thermal radiation

In the case of thermal radiation, energy is transferred in the form of electromagnetic waves. The mechanism of heat transfer from one body to another is not matter- or substance-related and is therefore also possible in a vacuum.

Figure 9: Changes of state of the Clausius-Rankine cycle

a) In the pV diagram, b) In the TS diagram.

1 Saturated-liquid line, 2 Saturated-vapor line, 3 Critical point.



Thermal conduction

Stationary one-dimensional thermal conduction in a simple flat wall

In a homogeneous body with constant cross-sectional area A the heat flow density $\dot{q} = \dot{Q}/A$ in the one-dimensional case is given by:

$$\dot{q} = -\lambda \frac{dT}{dx}.$$

This is the Fourier thermal-conduction equation with the (substance-specific) thermal conductivity λ of the wall. For the case shown in Figure 10 (expansion in the y - and z -directions very high) it follows for a temperature difference $T_1 - T_2$ between the walls:

$$\dot{q} = \frac{\lambda}{d} (T_1 - T_2)$$

and

$$\dot{Q} = \frac{\lambda}{d} (T_1 - T_2) A.$$

By analogy to the science of electricity, the thermal-conduction resistance can be introduced:

$$R_\lambda = \frac{T_1 - T_2}{\dot{Q}} = \frac{d}{\lambda A}.$$

Table 4 shows the values for λ for some materials.

Stationary one-dimensional thermal conduction in a layered flat wall

If the wall is made not of one but of several layered materials of different thicknesses $d_1, \dots, d_n$ and thermal conductivities $\lambda_1, \dots, \lambda_n$ (Figure 11), this produces for a temperature difference $T_1 - T_2$ normally for the layering:

$$\dot{q} = \frac{\lambda_R}{d} (T_1 - T_2),$$

with the effective resulting thermal conductivity:

$$\lambda_R = \frac{1}{\frac{d_1}{\lambda_1} + \frac{d_2}{\lambda_2} + \dots + \frac{d_n}{\lambda_n}}$$

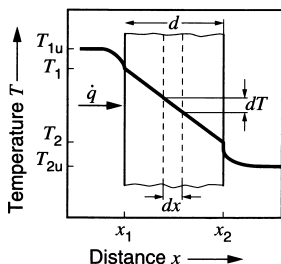
and the thermal resistance:

$$R_{\lambda R} = R_{\lambda 1} + R_{\lambda 2} + \dots R_{\lambda n}.$$

Table 4: Thermal conductivity λ of some materials at 293.15 K

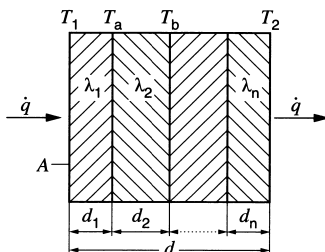
Material/medium	λ [W/K·m]
Aluminum (Al)	237
Iron (Fe)	81
Copper (Cu)	399
Titanium (Ti)	22
CrNi steel (X12CrNi 18.8)	15
Glass	0.87 – 1.40
Teflon (PTFE)	0.23
Polyvinyl chloride (PVC)	0.15

Figure 10: Stationary one-dimensional thermal conduction through a flat wall



SAND236-1E

Figure 11: Stationary one-dimensional thermal conduction through a layered, flat wall



SAND237Y

The thermal resistance therefore behaves additively (as in a series connection), with the order of the layers not playing a role.

If the temperature gradient is impressed along the plates (along a length l), it follows that:

$$\dot{q} = \frac{\lambda_p}{l} (T_1 - T_2)$$

with the effective resulting thermal conductivity

$$\lambda_p = \frac{1}{d} (\lambda_1 d_1 + \lambda_2 d_2 + \dots + \lambda_n d_n)$$

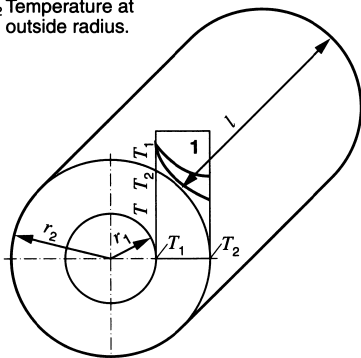
and the thermal-conduction resistance

$$\frac{1}{R_{\lambda p}} = \frac{1}{R_{\lambda 1}} + \frac{1}{R_{\lambda 2}} + \dots + \frac{1}{R_{\lambda n}}.$$

The reciprocal value of the total thermal resistance is therefore made up additively of the reciprocal values of the individual thermal resistances and behaves as in a parallel connection.

Figure 12: Stationary one-dimensional thermal conduction through a cylindrical tube wall

- 1 Temperature characteristic over the tube cross-section,
- r_1 Inside radius of tube,
- r_2 Outside radius of tube,
- l Length of tube,
- T_1 Temperature at inside radius,
- T_2 Temperature at outside radius.



SAND23B-3Y

Stationary one-dimensional thermal conduction in a tube wall

Similarly to thermal conduction through a wall, the following is obtained for a tube (of virtually infinite length) of length l in the one-dimensional case (thermal conduction only normally to the tube axis) as shown in Figure 12:

$$\dot{Q} = 2\pi l \lambda \frac{T_1 - T_2}{\ln \left(\frac{r_2}{r_1} \right)},$$

where T_1 is the temperature on the inner side and T_2 the temperature on the outer side.

Because of the different surface areas on the inner and outer sides of the tube the temperature assumes a logarithmic characteristic in the tube wall.

Stationary heat transmission

In many technical applications thermal conduction inside a solid body occurs in combination with a convective heat transfer (see Convective heat transfer). A (convective) heat transfer occurs as well as the thermal conduction, as shown in Figure 10; towards the wall on account of the temperature difference ($T_{1u} - T_1$) and away from the wall on account of the temperature difference ($T_2 - T_{2u}$). The heat flows must naturally all be of equal magnitude, since heat is not lost in either convective heat transfer or in thermal conduction. The heat flows that occur are governed by:

$$\dot{Q} = \frac{\lambda}{d} (T_1 - T_2) A,$$

$$\dot{Q} = \alpha_{1u} (T_{1u} - T_1) A,$$

$$\dot{Q} = \alpha_{2u} (T_2 - T_{2u}) A,$$

with the heat-transfer coefficients α_{1u} and α_{2u} . For such a situation it is possible to create the heat-transmission coefficient k , which is obtained as follows:

$$\frac{1}{k} = \frac{d}{\lambda} + \frac{1}{\alpha_{1u}} + \frac{1}{\alpha_{2u}}.$$

While taking into account the two fluid temperatures T_{1u} and T_{2u} , the following is obtained for the heat flow:

$$\dot{Q} = k (T_{1u} - T_{2u}) A.$$

Convective heat transfer

The heat-transfer coefficient α has already been introduced. It is dependent on a variety of influencing variables, such as, for example, temperature, density, geometry, type of flow (laminar or turbulent), and speed. It is determined by the conservation equation (nonlinear partial differential equations) describing the (fluidic) problem for which there are only closed analytical solutions in exceptional cases. Heat-transfer coefficients are therefore often determined on the basis of numerical solutions or directly from experiments. The functional dependencies for the heat-transfer coefficients can be found from these on the basis of dimensional analysis and similarity relationships.

In view of the complexity, reference is made at this point for derivation and further details to the relevant literature (e.g. [1]). The most important dimensionless parameter for heat transfer is the Nusselt number Nu :

$$Nu = \frac{\alpha L}{\lambda}.$$

This number contains, in addition to the heat-transfer coefficient α , the thermal conductivity λ of the fluid, and a reference length L , which can be for example a tube diameter or a plate length. It can further be shown that the following functional correlations apply to the averaged, i.e. integrated over the body surface and thus location-independent Nusselt number Nu_∞ :

- $Nu_\infty = f(Re, Pr)$ for forced convection,
- $Nu_\infty = f(Gr, Pr)$ for free convection.

The Nusselt number and thus the heat transfer can therefore be described as a function of the dimensionless parameters Reynolds number Re , Prandtl number Pr and Grashof number Gr . These must be calculated in each case as a function of the geometry and the flow (Re and Gr) or as a function of the fluid properties (Pr). The Reynolds number is defined as:

$$Re = \frac{cL}{\nu}.$$

Here c is the flow velocity, L a length characteristic of the application (e.g. tube diameter), and ν the kinematic viscosity of the fluid.

The Prandtl number is a pure substance variable and is given by:

$$Pr = \frac{\nu}{a},$$

with the thermal diffusivity a . For gases $Pr = 0.7$ can be assumed by approximation for pressures below 10 bar.

To calculate the Grashof number Gr , reference is made to the relevant literature (e.g. [1]).

The correlations for some selected cases with forced convection are set out in the following. The correlations in each case for the averaged values together with their range of application are specified here. The heat-transfer coefficients can be calculated with the Nusselt numbers determined from them and the definition of these.

Flat plate with laminar boundary layer subject to longitudinal flow
Nusselt correlation:

$$Nu_\infty = 0.664 Re^{\frac{1}{2}} Pr^{\frac{1}{3}}.$$

Characteristic length:
Plate length L .

Range of application:
 $Re \leq 5 \cdot 10^5$; $0.6 \leq Pr \leq 2,000$.

Flat plate with turbulent boundary layer subject to longitudinal flow
Nusselt correlation:

$$Nu_\infty = 0.037 Re^{0.8} Pr^{\frac{1}{3}}.$$

Characteristic length:
Plate length L .

Range of application:
 $5 \cdot 10^5 \leq Re \leq 10^7$; $0.6 \leq Pr \leq 60$.

Turbulent pipe inner flow
Nusselt correlation:

$$Nu_{\infty} = 0.023 Re^{0.8} Pr^n,$$

with $n = 0.4$ for $T_W > T_F$
and $n = 0.3$ for $T_W < T_F$.

Here, T_W is the wall temperature and T_F the averaged fluid temperature. The Nusselt number must be evaluated for the averaged fluid temperature because this changes during an inner flow from inlet to outlet.

Characteristic length:
Tube diameter D .

Range of application:
 $10^4 \leq Re$; $0.7 \leq Pr \leq 120$.

The tube length must be at least ten times the tube diameter.

The above correlation can also be used for tubes with non-circular cross-sections; the equivalent or hydraulic diameter d_h , which is created as follows, is used here:

$$d_h = \frac{4A}{U},$$

with the through-flow cross-sectional area A and the circumference U .

Thermal radiation

Thermal radiation depends only on the type and the temperature of the radiating body. When radiation strikes a body, the following phenomena are observed: Part of the incident radiation is reflected (reflection), part is absorbed (absorption), and part is let through (transmission). Due to the conservation of energy, it follows that the sum of these three energy components is equal to the amount of energy of the incident radiation.

According to the Stefan-Boltzmann law, the heat flow radiated from a body with surface A at temperature T is:

$$\dot{Q} = \varepsilon \sigma A T^4.$$

with the Stefan-Boltzmann constant σ ,

$$\sigma \approx 5.67 \cdot 10^{-8} \frac{\text{W}}{\text{m}^2 \text{K}^4},$$

and the emissivity ε of the body. The emissivity is between 0 (total reflection) and 1 (black-body radiator) and depends among others things on the temperature and the surface condition of the body. Some emissivity values are given by way of example in Table 5.

References

[1] H.D. Baehr, K. Stephan: Wärme- und Stoffübertragung. Springer, 9th Edition, Verlag Springer Vieweg, 2016.
[2] H.D. Baehr, S. Kabelac: Thermodynamik: Grundlagen und technische Anwendungen. 16th Edition, Verlag Springer Vieweg, 2016.
[3] E. Hahne: Technische Thermodynamik: Einführung und Anwendung. 5th Edition, Oldenbourg Wissenschaftsverlag, 2010.
[4] B. Weigand, J. Köhler, J. von Wolfersdorf: Thermodynamik kompakt. 4th Edition, Springer Vieweg, 2016.

Table 5: Emissivity ε of some materials (values in the range up to 300 °C)

Black-body radiator	1.00
Aluminum, unmachined	0.07
Aluminum, polished	0.04
Cast iron, rough, oxidized	0.94
Cast iron, turned	0.44
Copper, oxidized	0.64
Copper, polished	0.05
Brass, matt	0.22
Brass, polished	0.05
Steel, matt, oxidized	0.96
Steel, polished, oil-free	0.06
Steel, polished, oiled	0.40

Electrical engineering

Electromagnetic fields

Electrical engineering deals with electromagnetic fields and their effects. These fields are associated with electric charges (in each case an integral multiple of the electric elementary charge). Physics does not indicate whether the fields are the cause or the effect. Static charges produce an electric field, whereas moving charges give rise to a magnetic field as well. The association of electric and magnetic fields with static and moving charges is described by Maxwell's equations [1].

The presence of these fields is evidenced by the effects of their forces on other electric charges. The force on a point charge Q in an electric field is called Coulomb force. It causes a repulsion between charges of the same name. For two point charges Q_1 and Q_2 in free space at a distance a , it is:

$$F = \frac{Q_1 Q_2}{4\pi\epsilon_0 a^2}.$$

$\epsilon_0 \approx 8.854 \cdot 10^{-12}$ F/m is the electric field constant, also called the dielectric constant of free space.

The force acting on a moving charge in a magnetic field is expressed by the Lorentz force. This is responsible for two parallel conductors which carry rectified currents I_1 and I_2 being mutually attracting. Over a length l in free space at conductor distance a the force of attraction between the two conductors is:

$$F = \frac{\mu_0 I_1 I_2 l}{2\pi a}.$$

$\mu_0 \approx 1.257 \cdot 10^{-6}$ H/m stands for the magnetic field constant, also called the permeability of free space.

Electric field

The force acting on a static electric charge is ascribed to the effect of an electric field. An electrostatic field can be defined by the following quantities.

Electric potential $\varphi(P)$ and voltage U

The electric potential $\varphi(P)$ at point P is a measure of the work required per charge to move the charge Q from a reference point to point P:

$$\varphi(P) = \frac{W(P)}{Q}.$$

The voltage U is the potential difference (using the same reference point) between two points P_1 and P_2 :

$$U = \varphi(P_1) - \varphi(P_2).$$

Electric field strength E

The electric field strength E at point P depends on the location and its surrounding charges. It defines the maximum slope of the potential gradient at point P. The following applies to the field strength at a distance a from a positive point charge Q_1 : It is directed away from the charge Q_1 and has the value

$$E = \frac{Q_1}{4\pi\epsilon_0 a^2}.$$

Acting on a positive charge Q_2 at point P is a force in the direction of the electric field strength of the value

$$F = Q_2 E.$$

Electric field and matter

In a material which can be polarized (dielectric), an electric field generates electric dipoles (positive and negative charges $\pm Q$ at a distance a ; Qa is called the dipole moment). The dipole moment per unit volume is called the polarization M . The displacement density D indicates the density of the electric displacement flux, and is defined as follows:

$$D = \epsilon E = \epsilon_0 \epsilon_r E = \epsilon_0 E + M.$$

$\epsilon = \epsilon_0 \epsilon_r$ is the dielectric constant of the material, ϵ_0 the electric field constant (dielectric constant of vacuum), ϵ_r the permittivity (relative dielectric constant). For air, $\epsilon_r = 1$.

The quantity

$$w_e = \frac{1}{2} ED$$

is the electrical energy density. When multiplied by the volume, it produces the electrical energy W_e .

Table 1: Quantities and units

(Additional quantities and units in the text)

Quantity	SI unit
A Area	m^2
a Distance	m
B Magnetic flux density, induction	$T = Wb/m^2 = Vs/m^2$
C Electrical capacitance	$F = C/V$
D Electric flux density, electric displacement	C/m^2
E Electric field strength	V/m
F Force	N
f Frequency	Hz
G Electrical conductance	$S = 1/\Omega$
G Antenna gain	dB
H Magnetic H field strength	A/m
I Electric current	A
J Magnetic polarization	T
k Electrochemical equivalent	kg/C (usually: g/C)
L Inductance	$H = Wb/A = Vs/A$
l Length	m
M Electric polarization	C/m^2
P Power	$W = VA$
P_s Apparent power	VA
P_q Reactive power	$var = VA$
Q Quantity of electricity, electric charge	$C = As$
q Cross-sectional area	m^2
R Electrical resistance	$\Omega = V/A$
T Temperature	K
t Time	s
r Radius	m
r Reflectance factor	$-$
S Electromagnetic power density	W/m^2
s Standing wave ratio	$-$

Capacitor

Two metal bodies (electrodes) separated by a dielectric form a capacitor. When a voltage is applied to the capacitor, the two electrodes receive equal but opposite charges. The following equation holds for the received charge Q :

$$Q = CU.$$

C is the capacitance of the capacitor. It is dependent on the geometric shape of the electrodes, the distance by which they are separated, and the dielectric constant of the dielectric. Table 2 sets out the capacitances of selected arrangements.

Quantity	SI unit
U Voltage	V
V Magnetic potential difference	A
W Work, energy	$J = Ws$
W_e Electrical energy	Ws
W_m Magnetic energy	Ws
w_e Electrical energy density	Ws/m^3
w_m Magnetic energy density	Ws/m^3
w Number of turns	$-$
Z Characteristic impedance	Ω
α Geometric angle	$^\circ$ (degrees)
ϵ Dielectric constant	$F/m = C/(Vm)$
ϵ_0 Electric field constant	$\approx 8.854 \cdot 10^{-12} F/m$
ϵ_r Relative permittivity	$-$
λ Wavelength	m
Θ Current linkage	A
μ Permeability	$H/m = Vs/(Am)$
μ_0 Magnetic field constant	$\approx 1.257 \cdot 10^{-6} H/m$
μ_r Relative permeability	$-$
ρ Resistivity	$\Omega m = 10^6 \Omega mm^2/m$
σ Conductivity	$1/(\Omega m)$ $= 10^{-6} m/(\Omega mm^2)$
Φ Magnetic flux	$Wb = Vs$
ϕ Phase displacement angle	$^\circ$ (degrees)
$\phi(P)$ Potential at point P	V
ω Angular frequency	Hz $(= 2\pi f)$

Capacitance of capacitors connected in series and parallel:

$$\frac{1}{C_{\text{total}}} = \frac{1}{C_1} + \frac{1}{C_2}$$

(series connection, Figure 1a),

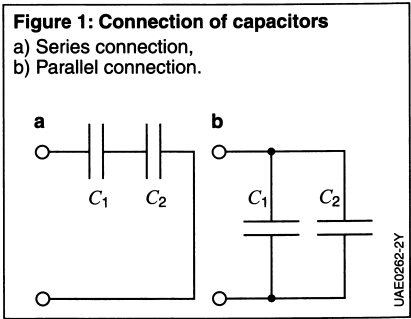
$$C_{\text{total}} = C_1 + C_2$$

(parallel connection, Figure 1b).

In the case of a parallel-plate capacitor (which also includes a wound capacitor), the inner plates of two capacitors connected in parallel act as electrodes (Figure 2).

The energy content of a charged capacitor (charge Q , voltage U , capacitance C) is:

$$W = \frac{1}{2}QU = \frac{Q^2}{2C} = \frac{1}{2}CU^2.$$



Direct current and direct voltage

Moving charges give rise to a current I , which is characterized by its intensity and measured in amperes. The direction of flow and magnitude of direct current are independent of time. The electrical system in a motor vehicle is, on the other hand, a direct-voltage system; its voltage is independent of time. The currents in a vehicle electrical system are usually time-dependent. In many cases time-dependent currents such as direct current can also be handled.

Current direction and measurement

Current flowing from positive pole to negative pole outside of the current source is designated as positive (in reality, the electrons travel from the negative to the positive

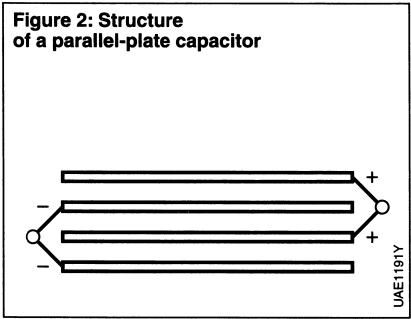


Table 2: Capacitance C of some conductor arrangements

Plate capacitor with n parallel plates	$C = (n - 1) \frac{\epsilon_r \epsilon_0 A}{a}$	ϵ_r, ϵ_0 Permittivity, electric field constant n Number of plates A Surface area of one plate a Distance between plates
Parallel conductors (twin conductors)	$C = \frac{\pi \epsilon_r \epsilon_0 l}{\ln \left(\frac{a + \sqrt{a^2 - 4r^2}}{2r} \right)}$	l Length of twin conductors a Distance between conductors r Conductor radius
Concentric conductor (cylindrical capacitor)	$C = \frac{2\pi \epsilon_r \epsilon_0 l}{\ln (r_2/r_1)}$	l Length of conductor r_2, r_1 Conductor radius where $r_2 > r_1$
Conductor to ground	$C = \frac{2\pi \epsilon_r \epsilon_0 l}{\ln \left(\frac{a + \sqrt{a^2 - r^2}}{r} \right)}$	l Length of conductor a Distance from conductor to ground r Conductor radius
Sphere with respect to distant surface	$C = 4\pi \epsilon_r \epsilon_0 r$	r Sphere radius

pole). Current measurement is performed by an ammeter (A) in the current path; voltage measurement by a voltmeter (V) connected in shunt (Figure 3).

The counting direction of voltage and current identified with the arrow can in principle be selected independently of the operating state. The accordingly oriented voltages and currents are thereby assigned a sign which is dependent on the operating state. In the case of positive voltages and currents, the arrow indicates the direction from positive (+) to negative (-).

Ohm's law

Ohm's law defines the relationship between voltage U and current I in solid and liquid conductors. The following applies:

$$U = RI.$$

The proportionality constant R is called ohmic resistance, and is measured in ohms (Ω). The reciprocal of resistance is called conductance G , and is measured in siemens (S).

$$G = \frac{1}{R}.$$

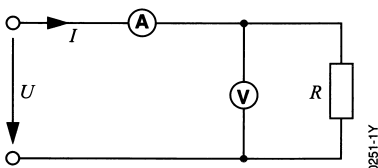
Ohmic resistance

Ohmic resistance depends on the material and its dimensions. For a solid wire:

$$R = \frac{\rho l}{q} = \frac{l}{q\sigma}.$$

Figure 3: Current and voltage measurement

R Load,
A Ammeter in current path,
V Shunt-connected voltmeter.



For a tube (from inside to outside):

$$R = \ln\left(\frac{r_2}{r_1}\right) \frac{1}{2\pi l\sigma}.$$

The equation's elements are:

ρ Resistivity [$\Omega \text{ mm}^2/\text{m}$],

$\sigma = 1/\rho$ Conductivity [$\text{m}/(\Omega \text{ mm}^2)$],

l Wire length, tube length [m],

q Wire cross-section [mm^2],

r_2, r_1 Outside and inside radii of tube [m]
with $r_2 > r_1$.

In the case of metals, resistance increases with temperature. The following applies:

$$R_T = R_{20} [1 + \alpha (T - 20^\circ\text{C})]$$

where

R_T Resistance at T ,

R_{20} Resistance at 20°C ,

α Temperature coefficient [$1/\text{K}$] ($= [1/^\circ\text{C}]$),

T Temperature [$^\circ\text{C}$].

Work and power

In a resistor through which current passes, the following holds for the energy converted during time t into heat or another form of energy (with ohmic resistance R , voltage U and current I):

$$W = UIt = RI^2 t$$

and thus for power:

$$P = UI = RI^2.$$

Kirchhoff's laws

First law: Current law

The sum of currents (according to their counting direction) flowing into each junction (node) is equal to the sum of currents flowing out of that junction.

Second law: Voltage law

For each closed loop of a conductor network the sum of the component voltages oriented in the direction of the loop at the individual elements (resistors and sources) is equal to the sum of the component voltages oriented against the direction of the loop.

Direct-current circuits

Circuit with load

In the circuit shown in Figure 4:

$$U = (R_a + R_l) I$$

where

R_a Ohmic resistance of load,

R_l Line resistance.

Battery-charging circuit

$$U - U_0 = (R_v + R_i) I \quad (\text{Figure 5})$$

Where

U Line voltage,

U_0 Open-circuit voltage (electromotive force) of battery,

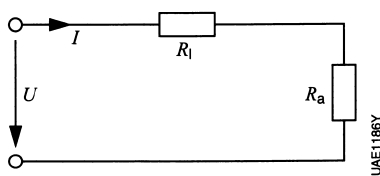
R_v Series resistance,

R_i Internal resistance of battery.

Condition for charging: $U > U_0$ (charging voltage greater than battery open-circuit voltage).

Figure 4: Circuit with load

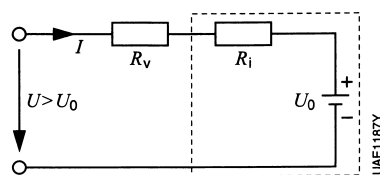
U Voltage,
 I Current,
 R_a Load resistance,
 R_l Line resistance.



UAE1186Y

Figure 5: Battery-charging circuit

U Line voltage,
 U_0 Open-circuit voltage of battery,
 R_v Series resistance,
 R_i Internal resistance of battery.



UAE1187Y

Series connection of resistors

$$R_{\text{total}} = R_1 + R_2 \quad (\text{Figure 6}),$$

$$U = U_1 + U_2.$$

The same current I flows in all the resistors (Kirchhoff's current law).

Parallel connection of resistors

$$\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} \quad \text{or} \quad G_{\text{total}} = G_1 + G_2;$$

$$I = I_1 + I_2; \quad \frac{I_1}{I_2} = \frac{R_2}{R_1} \quad (\text{Figure 7}).$$

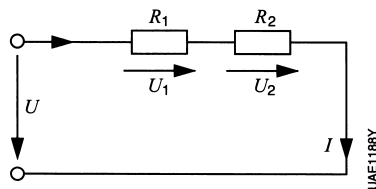
Voltage U is the same across all the resistors (Kirchhoff's voltage law).

Measurement of a resistance

A resistance can be measured by means of current and voltage measurement, and by using direct-reading ohmmeters or measuring bridges. Measuring bridges are used for example in pressure sensors to connect strain gages.

Figure 6: Series connection of resistors

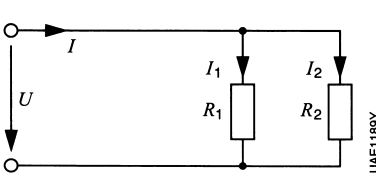
U Voltage,
 I Current,
 R Resistance.



UAE1188Y

Figure 7: Parallel connection of resistors

U Voltage,
 I Current,
 R Resistance.

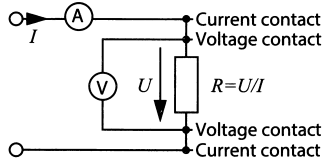


UAE1189Y

A wide used method of measurement, especially of very small resistances, is four-wire measurement (Figure 8). The measuring contacts for the test piece are configured in double form so as to prevent the resistances of the contacts from being measured as well. Only the small measured current flows through the voltage contacts to the voltmeter. The voltage drop caused by the large source current of the current source at the current contacts is no longer recorded.

Figure 8: Four-wire measurement of a resistance

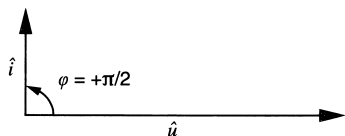
U Voltage,
 I Current,
 R Resistance.
 A Ammeter,
 V Voltmeter.



UAE1190-1E

Figure 9: Phasor diagram, capacitor

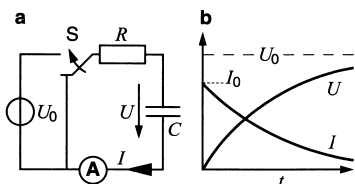
$\hat{u}$ Voltage amplitude,
 $\hat{i}$ Current amplitude.



UAE1193-2Y

Figure 10: Charging process at capacitor

a) Circuit,
b) Voltage and current curves.
 U Voltage,
 I Current,
 R Resistance.
 A Ammeter.



UAE1194-1Y

Time-dependent current

The relationships specified for direct current and direct voltage also apply in more or less modified form to time-dependent currents and voltages. In particular, Ohm's law and both of Kirchhoff's laws continue to apply unchanged.

Charging and discharging a capacitor

When a current I flows through a capacitor, its charge Q changes. The following applies:

$$I = \frac{dQ}{dt}.$$

In this way, there is a connection between current I and voltage U at the capacitor with capacitance C :

$$I = \frac{d(CU)}{dt}.$$

Here current I is supplied to that capacitor plate which has charge Q and from which voltage U is measured at the capacitor.

Examples

(C is constant in terms of time):

- Direct voltage U :
→ current $I = 0$.
- Initial voltage U_0 , direct current I :
→ voltage $U = U_0 + It$.
- Voltage harmonic (sinusoidal) with $U = \hat{u} \sin(\omega t)$:
→ current cosinusoidal with $I = \omega C \hat{u} \cos(\omega t) = \omega C \hat{u} \sin(\omega t + \frac{\pi}{2})$.

The quantity $\omega = 2\pi f$ is called the angular frequency in relation to frequency f , $\hat{u}$ is the voltage amplitude, and $\hat{i} = \omega C \hat{u}$ the current amplitude. The effective values $u_{\text{eff}} = \hat{u}/\sqrt{2}$ and $i_{\text{eff}} = \hat{i}/\sqrt{2}$ are also used instead of the amplitudes.

In the case of harmonic excitation, the current at the capacitor is therefore phase-displaced in relation to the voltage by the angle $\varphi = +\pi/2$ (it "leads"). This property is illustrated in a phasor diagram (Figure 9).

A further special case arises when the capacitor is charged via a resistance by a direct-voltage source U_0 (Figure 10) or discharged via a resistance. The time constant $\tau = RC$ is the decisive factor in the charging and discharging of a capacitor.

$$I = \frac{U_0}{R} e^{-\frac{t}{\tau}},$$

$$U=U_0(1-e^{-\frac{t}{\tau}}).$$

Discharging process

$$I = \frac{U_0}{R} e^{-\frac{t}{\tau}},$$

$$U = U_0 e^{-\frac{t}{\tau}},$$

U_0 : Charging voltage or voltage at start of discharge.

I Charging or discharging current,

$I_0 = U_0/R$ Current at start of charge,

R Charging or discharging resistance,

U Capacitor voltage.

Charging and discharging currents have opposite directions.

Magnetic fields are produced by moving electric charges, current-carrying conductors, magnetized bodies, or by an alternating electric field. They can be detected by their effect on moving electric charges (Lorentz force) or magnetic dipoles (like poles repel, and unlike poles attract).

Magnetic fields are characterized by the vector of the magnetic flux density B (induction). A straight conductor carrying current I_1 at distance a generates a magnetic flux density directed about it of the value

$$B = B_1 = \frac{\mu_0 I_1}{2\pi a}.$$

This magnetic flux density effects on a second conductor running in parallel with current I_2 equally directed over length l the force of attraction

$$F = B_1 I_2 l.$$

The magnetic flux density can be determined by means of voltage measurement, in that a changing magnetic field induces in a conductor loop a voltage:

$$U = \frac{d\Phi}{dt}$$

where

where
 $d\Phi$ change in the magnetic flux through
the conductor loop,
 dt change in time.

The magnetic flux density B is associated with the other field quantities as follows (q cross-section):

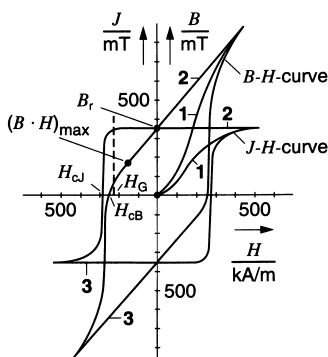
Magnetic flux $\Phi = Bq$.

In free space the following applies to magnetic field strength H :

$$H = \frac{B}{\mu_0}.$$

Figure 11: Hysteresis loop (e.g. hard ferrite)

- 1 Rise path,
2, 3 Demagnetization curves.
 H Magnetic field strength,
 B Magnetic flux density,
 J Magnetic polarization,
 B_r Remanence,
 H_{cB}, H_{cJ} Coercive field strength,
 H_G Limiting field strength.



Magnetic field and matter

In matter, induction B theoretically consists of two components. One component comes from the applied field ($\mu_0 H$), and the other from the matter (J) (see also the relationship between electric displacement density and electric field strength).

$$B = \mu_0 H + J$$

J is the magnetic polarization and describes that component of flux density contributed by the matter. In physical terms, J corresponds to one magnetic dipole moment per unit volume and is generally a function of field strength H . For many materials $J \gg \mu_0 H$ and is proportional to H . Thus:

$$B = \mu_r \mu_0 H$$

with relative permeability μ_r ; in free space it has the value $\mu_r = 1$.

The quantity

$$w_m = \frac{1}{2} B H$$

is called magnetic energy density. When multiplied by the volume, it produces the magnetic energy W_m .

Materials are divided into three groups according to their relative permeability values:

Diamagnetic materials

μ_r is independent of magnetic field strength and smaller than 1; the values are within the range:

$$(1 - 10^{-5}) < \mu_r < (1 - 10^{-11})$$

(e.g. Ag, Au, Cd, Cu, Hg, Pb, Zn, water, organic materials, gases).

Paramagnetic materials

μ_r is independent of magnetic field strength and greater than 1; the values are within the range:

$$(1 + 10^{-8}) < \mu_r < (1 + 10^{-4})$$

(e.g. O₂, Al, Pt, Ti).

Ferromagnetic materials

The magnetic polarization in these materials is very high, and its change as a function of the field strength H is nonlinear; it is also dependent on hysteresis. Nevertheless, if, as is usual in electrical engineering,

the relationship $B = \mu_r \mu_0 H$ is chosen, then μ_r is a function of H and exhibits hysteresis; the values for μ_r are within the range $10^2 < \mu_r < 5 \cdot 10^5$ (e.g. Fe, Co, Ni, ferrites).

Hysteresis loop

The hysteresis loop (Figure 11), which illustrates the relationship between B and H as well as J and H , is explained as follows: If the material is in the unmagnetized state ($B = J = 0$, $H = 0$) when a magnetic field H is applied, the magnetization follows the rise path (1). From a specific, material-dependent field strength, all magnetic dipoles are aligned and J reaches the value of saturation polarization J_s (material-dependent) which can no longer be increased. If H is now reduced, J decreases along section (2) of the curve and at $H = 0$ intersects the B or J axis at the remanence point B_r or J_r (in which case $B_r = J_r$). The flux density and polarization drop to zero only on application of an opposing field whose field strength is H_{CB} or H_{CJ} ; this field strength is called the coercive field strength.

As the field strength of the opposing field is further increased, saturation polarization is reached in the opposite direction. If the field strength is again reduced and the field reversed, curve (3), which is symmetrical to curve section (2), is traversed.

The following are usually tabulated as the most important characteristic values of a ferromagnetic material:

- Saturation polarization J_s
- Remanence B_r (residual induction for $H = 0$)
- Coercive field strength H_{CB} (demagnetizing field strength where B becomes equal to 0)
- Coercive field strength H_{CJ} (demagnetizing field strength where J becomes equal to 0, of significance only for permanent magnets)
- Limiting field strength H_G (a permanent magnet remains stable up to this field strength)
- Maximum small-signal permeability μ_{max} (maximum slope of the rise path; significant only for soft magnetic materials)
- Hysteresis loss (energy converted into heat per volume during one remagnetizing cycle, corresponds to the area of the B - H hysteresis loop; significant only for soft magnetic materials).

Ferromagnetic materials

Ferromagnetic materials are divided into soft and permanent magnetic materials. What must be emphasized is the immense range of eight powers of ten covered by the coercive field strength.

Permanent-magnet materials

Permanent-magnet materials have high coercive field strengths; the values lie within the range

$$H_{cJ} > 1 \frac{\text{kA}}{\text{m}}$$

Thus high demagnetizing fields H may occur without the material losing its magnetic polarization. The magnetic state and operating range of a permanent magnet lie within the 2nd quadrant of the hysteresis loop, on the demagnetization curve.

In practice, the operating point of a permanent magnet never coincides with the remanence point, because magnetic polarization of the permanent magnet in its interior always causes a magnetic field, the demagnetizing field, which shifts the operating point into the 2nd quadrant.

The point on the demagnetization curve at which the product BH reaches its maximum value $(BH)_{\text{max}}$ is a measure of the maximum attainable air-gap energy. In addition to remanence and coercive field strength, this value is important for characterizing permanent magnets.

AlNiCo, ferrite, FeNdB (REFe), and SeCo magnets are currently the most important types of permanent magnets in terms of industrial applications; their demagnetization curves (Figure 12) exhibit characteristics typical of the individual magnet types.

Soft magnetic materials

Soft magnetic materials have a low coercive field strength

$$H_{cJ} < 1 \frac{\text{kA}}{\text{m}},$$

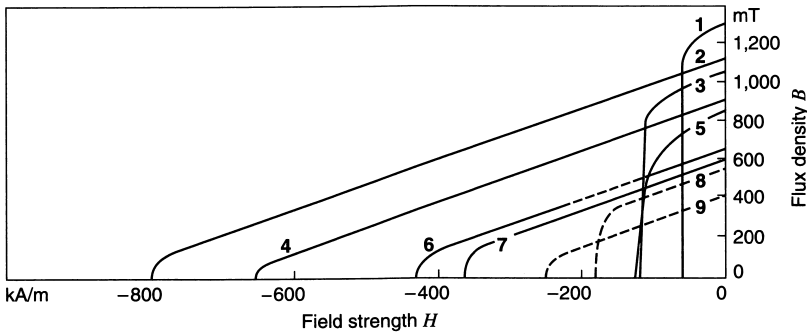
i.e. a narrow hysteresis loop. The flux density assumes high values (large μ , values) already for low field strengths so that, in customary applications, $J \gg \mu_0 H$, i.e. in practice no distinction need be made between $B(H)$ and $J(H)$ curves.

Due to their high induction at low field strengths, soft magnetic materials are used as conductors of magnetic flux. As they exhibit low remagnetization losses (hysteresis loss), materials with low coercive field strengths are particularly well-suited for applications in alternating magnetic fields.

The characteristics of soft magnetic materials depend essentially on their pretreatment. Machining increases the coercive field strength, i.e. the hysteresis loop becomes broader. The coercive field strength can be subsequently reduced to its initial value through material-specific annealing at high temperatures (magnetic

Figure 12: Demagnetization curves for various permanent-magnet materials

- 1 AlNiCo 52/6, 2 REFe 220/140, 3 AlNiCo 60/11, 4 SECo 112/100, 5 AlNiCo 30/10, 6 SECo 70/70p, 7 PiCo 60/40, 8 MnAl, 9 Hard ferrite 25/25.



final annealing). The magnetization curves, i.e. the B - H relationships, are set out in Figure 13 for several important soft magnetic materials.

Remagnetization losses

In Table 3 P1 and P1.5 represent the remagnetization loss for inductions of 1 and 1.5 tesla respectively, in a 50 Hz field at 20°C. These losses are composed of hysteresis losses and eddy-current losses. The eddy-current losses are caused by electric fields which are induced (law of

induction) in the magnetically soft components of the magnetic circuit as a result of changes in flux during alternating-field magnetization. Eddy-current losses can be kept low by applying the following measures to reduce electric conductivity:

- Lamination of the core
- Use of alloyed materials (e.g. silicon iron)
- Use of insulated powder particles (powdered cores) in the higher frequency range
- Use of ceramic materials (ferrites).

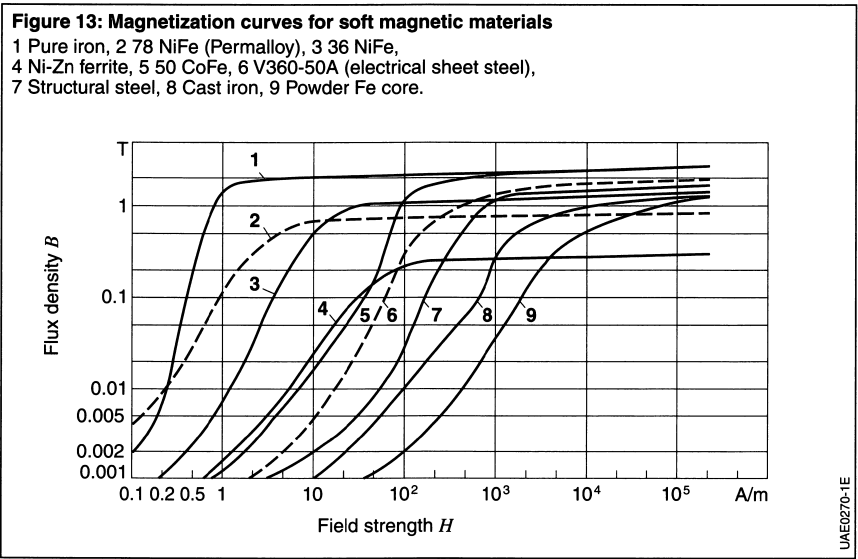


Table 3: Remagnetization losses

Sheet grade	Nominal thickness mm	Remagnetization loss W/kg		B (for $H = 10 \text{ kA/m}$) T
		P 1	P 1.5	
M 270 – 35 A	0.35	1.1	2.7	1.70
M 330 – 35 A	0.35	1.3	33.3	1.70
M 400 – 50 A	0.5	1.7	4.0	1.71
M 530 – 50 A	0.5	2.3	5.3	1.74
M 800 – 50 A	0.5	3.6	8.1	1.77

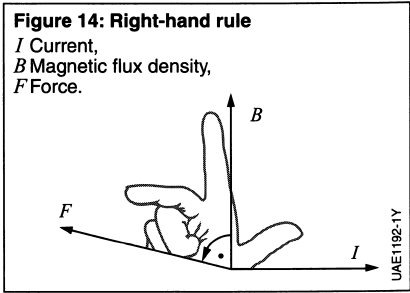
Magnetic field and electric current

Moving charges are associated with a magnetic field, i.e. conductors through which current flows are surrounded by a magnetic field. The direction in which the current flows (⊗ current flow into the page, ⊙ current flow out of the page) in the case of positive currents and the direction of the magnetic field strength form a right-handed screw. Table 4 shows the magnetic field strength of some conductor configurations.

In a magnetic field with flux density B a force is exerted on a current-carrying conductor (current I) of length l . If the conductor and the magnetic field form an angle of α , the following applies:

$F = B I l \sin \alpha$.

The direction of this force can be determined using the right-hand rule (Figure 14) (thumb pointed in the direction of current flow and the index finger in the direction of the magnetic field, then the middle finger points in the direction of force).



Law of induction

Any change in the magnetic flux Φ around which there is a conductor loop, caused for example by movement of the loop or changes in field strength, induces a voltage U_i at the terminals of the conductor

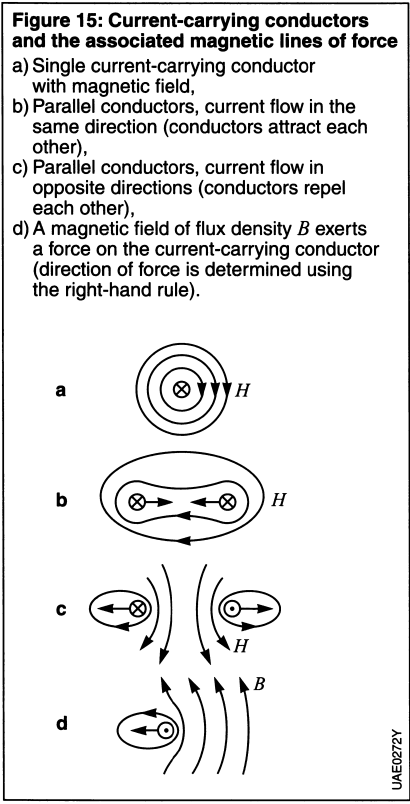


Table 4: Field strength of some conductor arrangements

Cylindrical coil, inner	$H = \frac{I w}{l}$	w I l	Number of turns Current through coil [A] Length of coil [m]
Straight wire in air, outside	$H = \frac{I}{2 \pi r}$	r I	Distance from wire axis [m] Current through coil [A]
Straight wire, inside	$H = \frac{I r}{2 \pi a^2}$	a r I	Wire radius [m] Distance from wire axis [m] Current through coil [A]
Circular wire, in center of circle	$H = \frac{I}{2 a}$	a I	Radius of circle [m] Current through wire [A]

loop. Thus, a voltage U_i is induced also in a (de-energized) conductor moving in direction v in the magnetic field (Figure 16):

$$U_i = Bvl$$

where

B Magnetic flux density,

l Conductor length,

v Velocity.

In a direct-current machine

$$U_i = 2 \frac{fz\Phi}{a},$$

where

U_i Induced voltage in V,

Φ Magnetic flux generated by the excitation (field) winding in Wb,

z Number of wires on the armature surface,

a Number of parallel armature-winding paths, frequency $f = pn/60$,

p Number of pole pairs,

n Rotational speed in rpm.

In an alternating-current machine

$$U_i = \frac{\pi f z \Phi}{\sqrt{2}}$$

where

U_i Effective value of induced voltage in V,

Φ Magnetic flux generated by the excitation winding or a permanent excitation in Wb,

$f = pn/60$ Frequency of alternating current in Hz,

p Number of pole pairs,

n Rotational speed in rpm,

z Number of wires on the armature surface.

In a transformer

$$U_1 = 2\pi f w_1 \Phi, \quad U_2 = 2\pi f w_2 \Phi$$

U_1, U_2 Effective values of induced voltages in V,

Φ Effective value of magnetic flux $\Phi(t)$ in Wb,

f Frequency in Hz,

w_1, w_2 Number of turns on the respective windings which surround the flux Φ .

The time-dependent flux $\Phi(t)$ is a superposition of contributions of the two time-dependent currents $i_1(t)$ and $i_2(t)$:

$$\Phi(t) = A_L(w_1 i_1(t) + w_2 i_2(t)).$$

A_L is called the AL value and is dependent on the transformer design.

The terminal voltage U is smaller (alternator) or larger (motor) than U_i by the ohmic drop in the winding (about 5%).

Self-induction

The magnetic field of a current-carrying conductor or a coil changes with the conductor current. A voltage proportional to the change in current is induced in the conductor itself. The following applies:

$$U = \frac{d(LI)}{dt}.$$

The inductance L depends on the relative permeability μ_r , which for most materials is practically equal to 1 and constant. Ferromagnetic materials are an exception. In the case of iron-core coils, therefore, L is highly dependent on the operating conditions. Table 5 shows the inductance L of some conductor configurations.

At low frequencies the inductance of the conductors is further increased by the inner inductance L_i of the wires. For round wires this is

$$L_i = \frac{\mu_0 l}{8\pi}.$$

Twin conductors of two round wires have double the value ($2 L_i$) per length l as inner inductance.

Figure 16: Motional induction

B Magnetic field,
 v Direction of moving conductor,
 U_i Positive induced voltage.

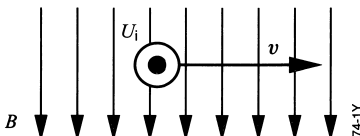


Table 5: Inductance of some conductor arrangements

Cylindrical coil	$L = w^2 \frac{\mu_r \mu_0 q}{l}$	μ_r, μ_0 Relative permeability, magnetic field constant w Number of turns on coil q Cross-section of coil [m ²] l Length of coil [m]
Parallel conductors (twin conductors) in air	$L = \frac{\mu_0 l}{\pi} \ln \frac{a + \sqrt{a^2 - 4r^2}}{2r}$	l Length of twin conductors [m] a Distance between conductors [m] r Conductor radius [m]
Concentric line (coaxial line)	$L = \frac{\mu_r \mu_0 l}{2\pi} \ln \frac{r_2}{r_1}$	l Length of line [m] r_2 Inside radius of outer tube [m] r_1 Radius of center tube [m]
Conductor to ground in air	$L = \frac{\mu_0 l}{2\pi} \ln \frac{a + \sqrt{a^2 - r^2}}{r}$	l Length of conductor [m] a Distance from conductor to ground [m] r Conductor radius [m]

Inductance of coils connected in series and parallel:

$$L_{\text{total}} = L_1 + L_2$$

(series connection, Figure 17a),

$$\frac{1}{L_{\text{total}}} = \frac{1}{L_1} + \frac{1}{L_2}$$

(parallel connection, Figure 17b).

The energy content of the coil carrying current I of inductance L is:

$$W = \frac{1}{2} L I^2.$$

Examples

(L is constant in terms of time):

- Direct current I :
→ voltage $U = 0$.
- Initial current I_0 , direct voltage U :
→ current $I = I_0 + U t$.
- Current sinusoidal with $I = \hat{i} \sin(\omega t)$:
→ voltage cosinusoidal with
 $U = \omega L \hat{i} \cos(\omega t)$,
 $U = \omega L \hat{i} \sin(\omega t + \pi/2)$.

The quantity $\omega = 2\pi f$ is called angular frequency in relation to frequency f ; $\hat{i}$ is the current amplitude, and $\hat{u} = \omega L \hat{i}$ the voltage amplitude. Here, too, the effective values $u_{\text{eff}} = \hat{u}/\sqrt{2}$ and $i_{\text{eff}} = \hat{i}/\sqrt{2}$ are frequently used.

In the case of harmonic excitation, the voltage at the coil is therefore phase-displaced in relation to the current by the angle $\varphi = +\pi/2$ (it "leads"). This property is illustrated in a phasor diagram (Figure 18).

A special case arises when the coil is connected via a resistance to a direct-voltage source U_0 or discharged via a resistance. The time constant $\tau = L/R$ is the decisive factor in the switch-on and switch-off operations of the coil.

Figure 17: Connection of coils

- a) Series connection,
b) Parallel connection.

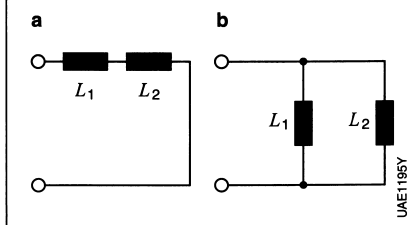
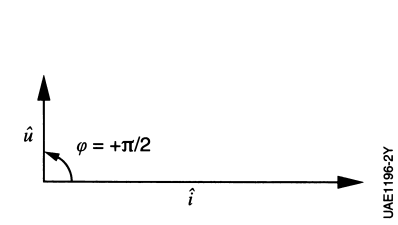


Figure 18: Phasor diagram, coil

- $\hat{u}$ Voltage amplitude,
 $\hat{i}$ Current amplitude.



Switch-on operation (Figure 19):

$$U = U_0 e^{-\frac{t}{\tau}}.$$

$$I = \frac{U_0}{R} (1 - e^{-\frac{t}{\tau}}).$$

Switch-off operation:

$$I = I_0 e^{-\frac{t}{\tau}}.$$

$$U = I_0 R e^{-\frac{t}{\tau}}.$$

Where

U_0 Excitation voltage,

I Coil current,

I_0 Coil current at start of switch-off operation,

R Resistance in series to coil,

U Coil voltage.

The coil currents during the switch-on and switch-off operations have opposite directions.

The magnetic circuit

In addition to material equations, the following equations also determine the design of magnetic circuits:

1. Ampère's law

(equation of magnetic voltage)

The following equation holds true for a closed magnetic circuit:

$$\sum H_i l_i = \sum V_i = I w,$$

where $I w = 0$ if no current is surrounded by the magnetic circuit.

$I w = \Theta$ is the magnetomotive force (ampere turns, current surrounded in total by the magnetic circuit),

$H_i l_i = V_i$ magnetic potential difference ($H_i l_i$ is to be calculated for circuit components in which H_i is constant).

2. Law of continuity

(equation of magnetic flux).

The same magnetic total flux $\Phi = B A$ flows in the individual sections of a magnetic circuit.

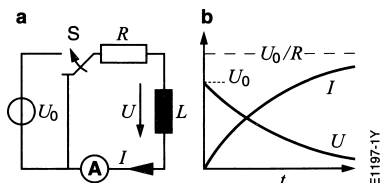
$\Phi = \text{const.}$ in all sections, A is the cross-section of the respective section.

Within a section of the magnetic circuit the magnetic flux can split into partial fluxes $\Phi_1, \Phi_2, \dots$ the sum of which however is again the constant total flux Φ .

The quality of a circuit is determined by the amount of flux available in the working air gap. This flux is called useful flux. The leakage flux – the difference between the total flux and the useful flux – does not pass through the working air gap and does not add to the power of the magnetic circuit. Its ratio to total flux (flux of the permanent magnet or electromagnet) is called the leakage coefficient σ (practical values for σ are between 0.2 and 0.9).

Figure 19: Switch-on operation for coil

- a) Circuit,
b) Voltage and current curves.



Wave propagation

Waveguide

At higher frequencies forward and return conductors of an electric connection section are combined and called a waveguide or in short a line. Common structures are the parallel-wire line, the twisted two-wire line (twisted pair), the coaxial line, the strip line, and the microstrip line.

At higher frequencies the current is no longer distributed evenly over the line cross-section. In the individual wire the current is increasingly displaced to the edge as the frequency increases (skin effect). In lines there is an additional displacement of current to the other conductor in each case (proximity effect). In particular the ohmic loss resistance and thus the damping of the line are thereby increased.

A line has the characteristic impedance Z if it is terminated with this resistance Z and then demonstrates the same impedance Z as the input resistance. In the lines used in motor vehicles (TEM or L lines) the characteristic impedance Z is directly connected with the capacitance C per length l and the inductance L per length l :

$$Z = \sqrt{\frac{L}{C}}.$$

Table 6 shows characteristic impedances for some selected line types.

Even the amplitudes of the electric and magnetic fields along the line are linked via the characteristic impedance:

$$Z = \frac{\hat{E}}{\hat{H}}.$$

A line of characteristic impedance Z which is terminated with resistance R demonstrates at its end the reflectance factor

$$r = \frac{R - Z}{R + Z}.$$

The square of the absolute value of the reflectance factor indicates what fraction of the power is sent back at the end of the line.

$$P_r = |r|^2 P_0.$$

Passed on to the terminating resistor, however, the power becomes

$$P_t = (1 - |r|^2) P_0.$$

When $R = Z$ applies, adaptation occurs and the reflectance factor r is zero. In this case, no power is reflected, the full power being passed on to the terminating resistor R .

At higher frequencies and with longer lines, voltage and current along the line are no longer constant. At a fixed instant both vary along the line. Here there can be points over the length of the line at which the amplitude of the voltage fluctuations is at maximum, while at other points only minimum voltage amplitudes occur. The ratio of maximum to minimum amplitude on a line is called the standing wave ratio s .

$$s = \frac{U_{\max}}{U_{\min}}.$$

If a line is supplied at one end only with a sinusoidal time function, while terminated at the other end with a resistor, it is possible to calculate the standing wave ratio from the reflectance factor.

$$s = \frac{1 + |r|}{1 - |r|}.$$

Table 6: Characteristic impedances for some selected line types

Parallel conductors (twin conductors) in air	$Z = \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{1}{\pi} \ln \frac{a + \sqrt{a^2 - 4r^2}}{2r}$	a r	Distance between conductors [m] Conductor radius [m]
Concentric line (coaxial line)	$Z = \sqrt{\frac{\mu_r \mu_0}{\epsilon_r \epsilon_0}} \frac{1}{2\pi} \ln \frac{r_2}{r_1}$	r_2 , r_1	Inside radius of outer tube [m] Radius of center tube [m]
Conductor to ground in air	$Z = \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{1}{2\pi} \ln \frac{a + \sqrt{a^2 - r^2}}{r}$	a r	Distance from conductor to ground [m] Conductor radius [m]

During adaptation $s = 1$ and at all points on the line the voltage amplitude is the same.

Couplers and matching transformers

Couplers are used in conjunction with matching transformers to forward an electromagnetic wave from one waveguide to another. Matching transformers reduce the reflectance factor during the transition from one characteristic impedance to another or couple a symmetrical waveguide (e.g. a parallel-wire line) with an asymmetrical one (e.g. a coaxial line).

Couplers and matching transformers are usually reciprocal, coupling just as well in one direction as in the other.

Wave propagation in free space

Even free space is a waveguide, with the characteristic impedance of free space (field characteristic impedance Z_0):

$$Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} \approx 377 \, \Omega.$$

Far away from antenna structures the direction of the electric field, the direction of the magnetic field and the direction of propagation of the electromagnetic wave in free space are vertically on top of one another, and the electric field strength is linked with the magnetic field strength via the characteristic impedance Z_0 . The electromagnetic power density S emitted by a transmitter is established via the amplitudes of the electric field strength $\hat{E}$ and the magnetic field strength $\hat{H}$. In free space, far away from the transmitter, the electromagnetic power density decreases with the inverse square of the distance from the transmitter:

$$S = \frac{1}{2} \hat{E} \hat{H};$$

$$S(r) = S(r_0) \frac{r_0^2}{r^2}.$$

Electric and magnetic field strengths, on the other hand, decrease only inversely with the distance from the transmitter:

$$\hat{E}(r) = \hat{E}(r_0) \frac{r_0}{r},$$

$$\hat{H}(r) = \hat{H}(r_0) \frac{r_0}{r}.$$

For some applications in motor vehicles the receiver is situated so near to the transmitter that near-field conditions prevail. Radio remote controls, for example, generate via a current-carrying coil magnetic fields whose field strength in the near field is inversely proportional to the cube of the distance from the transmitter.

Antennas

A coupler from a waveguide to free space is called an antenna. Antennas too are usually reciprocal; they can therefore act equally as transmitting and receiving antennas.

An electromagnetic wave in free space consists of time-variant electric and magnetic fields. These have one spatial direction – far away from transmitting antennas the field directions of these two fields are vertical to the direction of propagation of the wave. The direction of the corresponding electric field determines the polarization direction of the electromagnetic wave. Most antennas are designed in such a way that they can transmit and receive only one single direction of the electric (sometimes also the magnetic) field strength. They therefore transmit only linearly polarized waves and receive only one particular polarization of the wave.

A distinction is made between narrow-band antennas (dipole antennas, as used for example for GPS, GSM, Bluetooth, WLAN or even VHF) and wide-band antennas, which are used as transmitting and receiving antennas in EMC measurements in motor vehicles. An example of a narrow-band antenna is the half-wave dipole, a rod or wire of length $l = 0.96 \lambda/2$ (λ : wavelength in free space). The optimum length becomes shorter for thicker rods, or if necessary also through connection with additional capacitances. In conjunction with an extended metal surface the dipole length is shortened by half to one $\lambda/4$ -dipole.

An example of a wide-band antenna is the logarithmically periodic antenna. Here, several half-wave dipoles are of different lengths are arranged along a twin line. No antenna transmits in transmitting mode uniformly in all directions. Likewise, no antenna receives in receiving mode all directions equally well. This directionality is described by the directional diagram (also known as the antenna directional characteristic). Figure 20 shows the directional diagram of a half-wave dipole in a vertical section through the dipole antenna. What is depicted is the radiated power density against the angle.

Often what is of interest is not the entire directional diagram, but only its maximum value in the preferred direction. The directivity indicates by how much the radiated power per surface area $S_{\max}$ of the antenna in transmitting mode is greater in the preferred direction than for a reference antenna which does not prefer a direction:

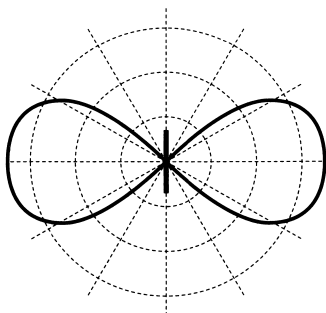
$$D_0 = \frac{S_{\max}}{S_0}.$$

The directivity in relation to the antenna gain G is given as a dB value:

$$G = 10 \log D_0.$$

Figure 20: Directional diagram of a half-wave dipole

(Vertical section through the dipole antenna, power density against the angle).



UAE1198Y

The antenna losses are usually also included in the antenna gain such that G is smaller than in the ideal case. The same directivity and the same antenna gain also apply in receiving mode. There they indicate by how much the antenna power input is greater than for the reference antenna.

In the EMC standards calculations are made not with the antenna gain but with the antenna factor AF :

$$AF = 20 \log \frac{E}{U}.$$

The antenna factor indicates what voltage amplitude U (in V) drops at the input resistance of the measuring device connected to the antenna when an electric field of amplitude E (in V/m) is received. The same antenna factor also applies to transmitting mode and links the voltage amplitude U (in V) at the input of the transmitting antenna with the amplitude E (in V/m) of the electric field of the emitted wave.

Antenna gain and antenna factor can be converted into one another:

$$AF = 10 \log \left(\frac{4\pi Z_0}{\lambda^2 R_L} \right) - G$$

where

R_L Input resistance of measuring device and of antenna input (in Ω , usually 50 Ω),
 Z_0 Field characteristic impedance of free space (equal to 377 Ω),
 λ Wavelength (in m).

Electric effects in metallic conductors

Contact potential between conductors

Contact potential occurs in conductors, and is analogous to the triboelectricity or voltaic emf in insulators (e.g. glass, hard rubber). If two dissimilar metals at the same temperature are joined to make metal-to-metal contact with one another and are then separated, a contact potential is present between them. This is caused by the different work functions of the electrons. The magnitude of contact potential depends on the element positions in the thermoelectric series (Table 7). If more than two conductors are so joined, the resulting contact potential is the sum of the individual contact potential values.

Thermoelectricity and thermocouples

Seebeck effect

A voltage occurs between the two ends of a conductor if these exhibit different temperatures. As the temperature increases the electron density in the conductor decreases on account of the increasing kinetic energy of the electrons: The electrons in the "hot" zone have a greater kinetic energy and move on average towards the "cold" zone. In all, there is then a higher electron density in the "cold" zone than in the "hot" zone and hence a negative potential. In semiconductor physics this can be expressed by different Fermi levels and a resulting potential difference (thermoelectromotive

force U_T). In all, the thermoelectromotive force U_T between the ends is obtained with the temperature difference ΔT according to the following equation:

$$U_T = \alpha \Delta T,$$

where

α : Seebeck coefficient in $\mu\text{V/K}$,
 ΔT : Temperature difference in K.

The Seebeck coefficient α is approximately temperature-independent and is determined by the material. The Seebeck coefficient α can be used to determine the thermoelectromotive force U_T in μV for the respective material for a given temperature difference in kelvin; frequently the second conductor is made of platinum – therefore the Seebeck coefficient α for platinum in Table 8 equals zero. Occasionally the k factor is used instead of α .

Technically, this effect cannot be used with a homogeneous material. Since the two measuring points are brought to the same temperature for the voltage tap via

Table 7: Placement of some elements in the thermoelectric series at the temperature 0 °C (source [5])

(For Pb the thermoelectromotive force is arbitrarily set at zero)

Element	Thermoelectromotive force [$\mu\text{V K}^{-1}$]
Antimony	+35
Iron	+16
Zinc	+3
Copper	+2.8
Silver	+2.7
Lead	0
Aluminum	-0.5
Platinum	-3.1
Nickel	-19
Bismuth	-70

Table 8: Seebeck coefficient of some elements referred to platinum at the standard temperature 0 °C (source [6])

Element	Thermoelectromotive force [$\mu\text{V K}^{-1}$]
Selenium	900
Tellurium	500
Silicon	440
Germanium	300
Antimony	47
Nickel chromium	25
Iron	19
Tungsten	7.5
Cadmium	7.5
Silver	6.5
Gold	6.5
Copper	6.5
Rhodium	6
Tantalum	4.5
Lead	4
Aluminum	3.5
Carbon	3
Mercury	0.6
Platinum	0
Sodium	-2
Potassium	-9
Nickel	-15
Constantan	-35
Bismuth	-72

the measuring device, two equal voltages would be connected in series in the opposite direction. Consequently, no voltage would be tapped between the two measuring points. This changes when different materials are used (Figure 21). The actual thermocouple consists in Figure 21 of two conductors made of materials 1 and 2. Since the connections are in some cases too short, they are extended at a coupling point. This coupling point at the temperature T_K . If necessary, a further extension is required – this is depicted in Figure 21 with material 3. This can involve, for example, conductor tracks of copper. Finally, the thermoelectromotive force U_T can be tapped. The electronic circuit here is at the ambient temperature T_U .

The series connection of the different elements in Figure 21 produces the following relationship:

$$U_T = \alpha_3 (T_U - T_K) + \alpha_1 (T_K - T_M) + \alpha_2 (T_M - T_K) + \alpha_3 (T_K - T_U),$$

$$U_T = (\alpha_1 - \alpha_2) (T_K - T_M).$$

To match the material pairing of a thermocouple corresponding extensions of the same material can be used as compensating lines.

The advantage of thermocouples is their small size, enabling them to be used in areas with very small dimensions – for example, thermocouples can be stuck to the

surface for flow measurements in a wind tunnel. The thermoelectromotive forces are indeed very low; however, the internal resistance of the thermocouples is low such that disturbances have little bearing on the output signal. Thermocouples can be used to good effect up to temperatures in excess of 2,000 °C. On the down side, however, thermocouples are subject to low long-term stability.

Peltier effect

The reciprocal of the Seebeck effect is the Peltier effect, in which a temperature difference is created by electrical energy (e.g. as a heat pump). A current flow through the contact points of two different conductors results in heat energy Q being absorbed at one contact and picked off at the other. Depending on the current direction, heat can thus be supplied or dissipated at a contact. The heat transfer ΔQ per time unit Δt is governed by the following equation:

$$\frac{\Delta Q}{\Delta t} = \Pi I \text{ (in VA)},$$

where

Π : Peltier coefficient (in V),

I : Current intensity (in A),

Δt : Time interval (in s).

The relationship between the Peltier coefficient Π , the temperature T and the thermoelectromotive force U_T is as follows:

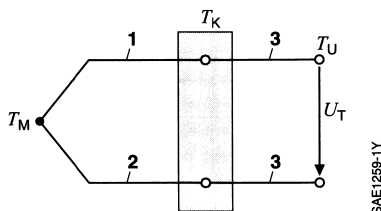
$$\Pi = U_T T.$$

To utilize the Peltier effect, bismuth telluride (Bi_2Te_3) or even silicon-germanium (SiGe) for example is used; these are either n - or p -doped so as to obtain zones with different Fermi levels. The maximum voltage per thermocouple (i.e. per n - and p -doped zone) is 0.12 V so that a Peltier element with 127 thermocouples requires a supply voltage of approximately 15 V.

It must be noted when operating a Peltier element that the cooling capacity does indeed increase with increasing current I , but this results on account of the electric (ohmic) power loss $P = RI^2$ in self-heating, which counteracts the cooling capacity. Consequently, determining the operating point with maximum efficiency is governed by the following: The ratio between current

Figure 21: Basic design of a thermocouple

- 1 Conductor with material 1,
- 2 Conductor with material 2,
- 3 Conductor extension with material 3,
- T_M Temperature at the measuring point,
- T_K Temperature at the coupling point,
- T_U Ambient temperature,
- U_T Thermoelectromotive force.



intensity and maximum current intensity must be chosen in the same way as that between the temperature difference and the maximum achievable temperature difference.

Galvanomagnetic and thermomagnetic effects

Such effects are understood to be changes caused by a magnetic field in the flow of electricity or heat within a conductor. There are twelve different recognized effects which fall into this category, the most well-known of which are the Hall, Ettingshausen, Righi-Leduc, and Nernst effects.

Hall effect

The Hall effect is of particular significance in technical applications. If a current I is sent through a suitable conductor and simultaneously a magnetic field or a magnetic flux density B is applied perpendicular to this current, a voltage is produced which is perpendicular to both the flow of current and the magnetic field. This voltage is called the Hall voltage U_H (Figure 22):

$$U_H = \frac{R_H I B}{d},$$

where

R_H Hall constant (in m^3/As),

I Supply current (in A),

B Magnetic flux density (in Vs/m^2),

d Thickness of conductor (in m).

In the case of Hall-effect sensors made of doped materials such as silicon – i.e.

n - or p -type silicon – the majorities determine the sign and the direction of the Hall voltage U_H .

The Hall constant R_H is relevant to use: Highest values in the range of $10^{-4} \text{ m}^3/\text{As}$ are achieved for indium arsenide (InAs); for silicon, the values are among other things dependent on the doping and are much lower. When the current I is known, it is possible to infer the magnetic flux density B or the magnetic field. The advantage here is the strong linear dependence of the Hall voltage on the magnetic field or the magnetic flux density; the disadvantage is as a rule the not insignificant offset on account of inhomogeneities (crystal defects, particularly on the surface) and geometry influences.

In ferromagnetic materials, the Hall voltage is a function of magnetization, where hysteresis effects can also occur.

Hall-effect sensors are frequently used for rotational-speed recording, in which the alternation of tooth and gap of a rotating soft-magnetic sensor wheel is detected. Other applications in automobiles include the detection of end points or the potential-free (i.e. contactless) measurement of electric currents via the magnetic field generated by the current.

Magnetoresistive effects

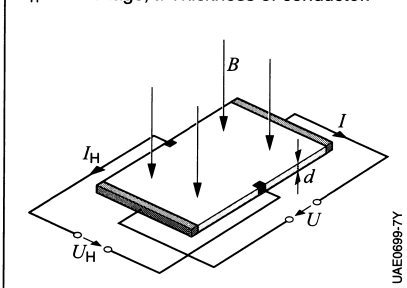
Anisotropic magnetoresistive effect

Sensors based on the anisotropic magnetoresistive effect (AMR effect) change the electrical resistance R on account of an applied magnetic field H ; in contrast to conventional Hall-effect sensors and also magnetoresistors, AMR sensors react to magnetic fields in the layer plane (here xy plane in Figure 23).

AMR sensor elements essentially consist of a thin strip roughly only 20 to 50 nm thick of highly permeable permalloy ($\text{Ni}_{81}\text{Fe}_{19}$ alloy), which exhibits strong shape anisotropy (Figure 23). This is expressed by a much greater length l than width b ; this defines an “easy axis”, which determines the orientation of magnetization M without an external magnetic field H in the xy plane in Figure 23. A magnetic field H in the direction of the “hard axis”, in Figure 23 in the y direction) rotates M out of this position, which results in a relative change of the electrical resistance R by max. 3 %.

Figure 22: Hall-effect sensor element

B Flux density, I Supply current,
 U Supply voltage, I_H Hall current,
 U_H Hall voltage, d Thickness of conductor.



R depends on the angle Θ_{JM} between the current density J and the magnetization M : If J and M are parallel (i.e. $\Theta_{JM}=0$), the maximum electrical resistance $R_{||}$ results; if J and M are perpendicular to each other ($\Theta_{JM}=90^\circ$), the minimum value $R_{\perp}$ is obtained. The term $\sin \Theta_{JM}$ can be expressed for small angles Θ_{JM} by the ratio of the applied magnetic field H_y in the hard direction and the anisotropy field strength H_K , where H_K depends on the geometry of the AMR

layer – more precisely on the ratio of width b to thickness d . The following relationship applies to R (Figure 23b, [7]):

$$R(\Theta_{JM}) = R_{||} - \Delta R_{\max} (\sin \Theta_{JM})^2.$$

The following applies to small Θ_{JM} :

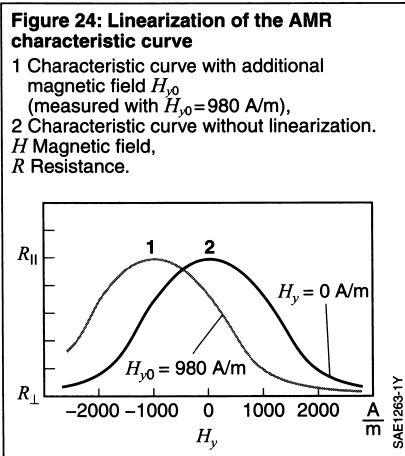
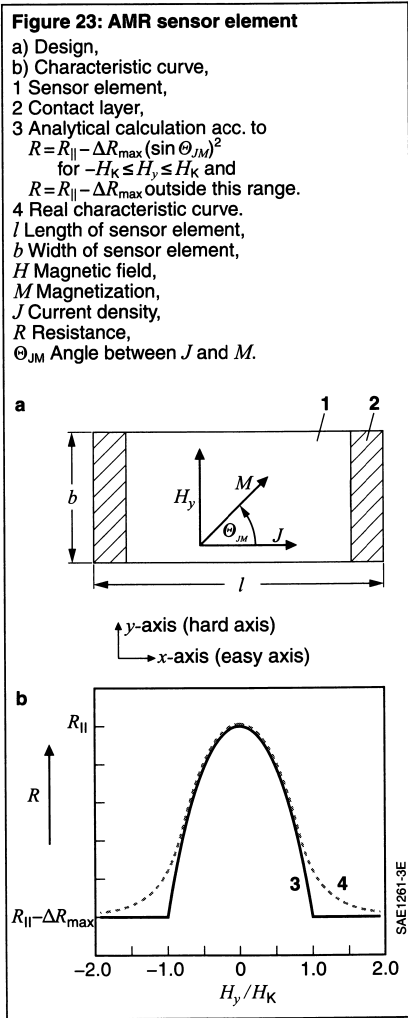
$$\sin \Theta_{JM} \approx \frac{H_y}{H_K}.$$

- $R_{||}$: Resistance for $\Theta_{JM} = 0^\circ$;
 $R_{\perp}$: Resistance for $\Theta_{JM} = 90^\circ$;
 $\Delta R_{\max} = R_{||} - R_{\perp}$: Maximum achievable absolute resistance change.

Quantum physics provides a physical explanation of this effect: Changes in the electric conductivity are obtained depending on the spin orientation.

The anisotropy field strength H_K plays an important role for the sensitivity $S = dR/dH$; the wider the AMR layer, the steeper the characteristic curve and hence the sensitivity S .

It follows from Figure 23b that for small magnetic fields H_y , the sensitivity S is very low; the operating point should therefore be brought into a zone with linear behavior. Above all two methods are known for this purpose: Linearization with an additional (constant) magnetic field $H_{y,0}$ (in the direction of the hard axis in Figure 23a), which shifts the characteristic curve in Figure 24 to the left into the steep zone;

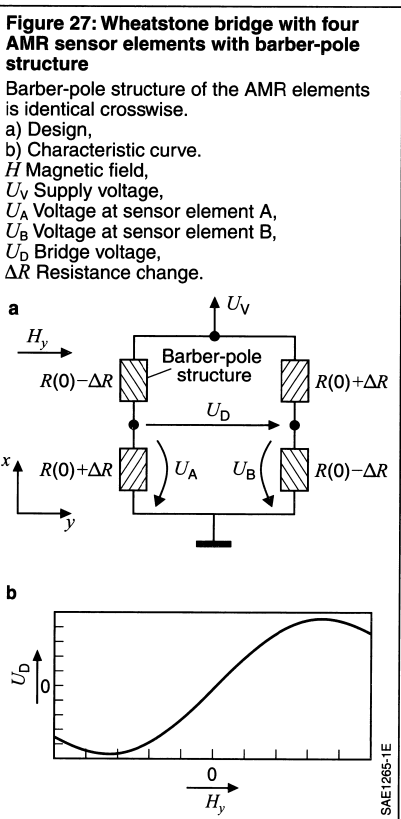
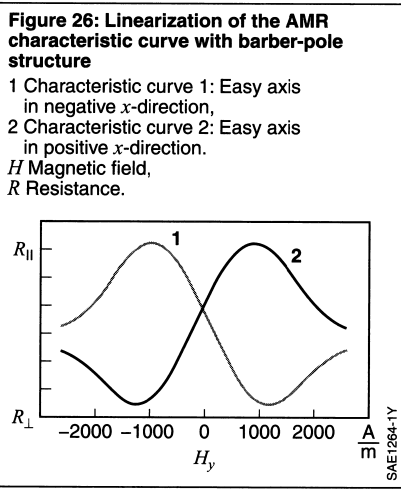
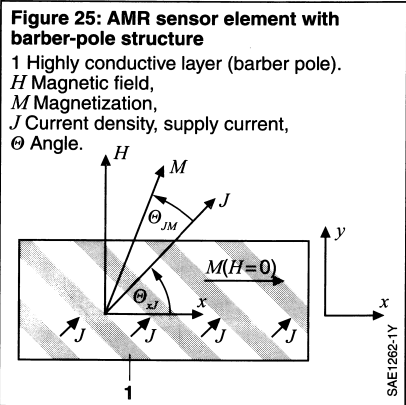


or the frequently used barber-pole structure in Figure 25. Here the current-density vector J is rotated through approximately 45° to obtain a linear characteristic curve for small magnetic fields H_y .

However, it is to be noted according to Figure 26 that the orientation of the easy axis may not be changed since "rotating the domains around an axis" through 180° leads to a mirroring of the characteristic curve with regard to $H_y=0$ ("butterfly curve"). This is as a rule effected by an additional constant magnetic field in the "easy direction" (here: x -direction), but which reduces the sensitivity and also the absolute resistance and hence also the maximum relative change of R is always approximately 3 %. Rotating the domains around an axis results in a change of the characteristic curve during linearization with a constant magnetic field H_{y0} .

The use of AMR sensors is favored by the higher magnetic-field sensitivity obtained than with silicon Hall-effect sensors. On the down side, AMR sensors are not compatible with semiconductor processes and must therefore be designed separately. Because of the mirror-symmetric characteristic curve around zero AMR angle sensors can only resolve an angular range of 180° .

Figure 27 shows a typical circuit of four AMR elements with barber-pole structure; because of the opposing orientation of the highly conductive strips in two AMR



elements respectively a piecewise linear characteristic curve is achieved, as can be seen by reference to Figure 27b.

Giant magnetoresistive effect

In the case of the giant magnetoresistive effect (GMR effect) the electrical resistance R changes with the orientation of the magnetization M of two closely neighboring ferromagnetic layers (Figure 28, e.g. iron or cobalt) which are separated by a very thin non-magnetic intermediate layer (e.g. chrome). As with the AMR effect, only applied magnetic fields in the layer plane are relevant with regard to the GMR effect. The spin-dependent scattering of the electrons in one of the magnetic layers brings about, with a parallel orientation of the magnetization M (i.e. ferromagnetic coupling), the minimum electrical resistance $R_{\min}$ in the two layers; with an antiparallel orientation of the two magnetizations M (thus

antiferromagnetic coupling) the maximum value $R_{\max}$ is the result. Relative resistance changes of approximately 6 % are obtained when a number of layer sequences are used.

The current I flows in GMR sensors preferably in the layer plane of the magnetization M (CIP-GMR, Current-in-plane-GMR). However, a current flow perpendicular to the layer plane (CPP-GMR, Current-perpendicular-to-plane-GMR) is also possible. Figure 28b shows the bell shape of R (similar to the AMR effect) as a function of an applied magnetic field H .

The important aspect here is the minimal thickness of the electrically conductive and non-magnetic intermediate layer (chrome in Figure 28 or "spacer" in Figure 29; copper is also used), which is in the region of below one nanometer.

GMR sensor elements are frequently used as spin valves in order to generate a sufficient and at least piecewise linear signal in the case of small magnetic fields H . The basic design is shown in Figure 29: Here the magnetization M_2 of a ferromagnetic layer is retained (in the "pinned layer" with antiferromagnetic coupling) so that only the ("rotatable") magnetization M_1 of the soft-magnetic layer follows an external magnetic field H_y ; M_x in Figure 29 corre-

Figure 28: Characteristic curve and basic design of a GMR sensor element

- a) Basic design,
b) Characteristic curve.
 M Magnetization,
 I Current,
 R Resistance,
 H Magnetic field.

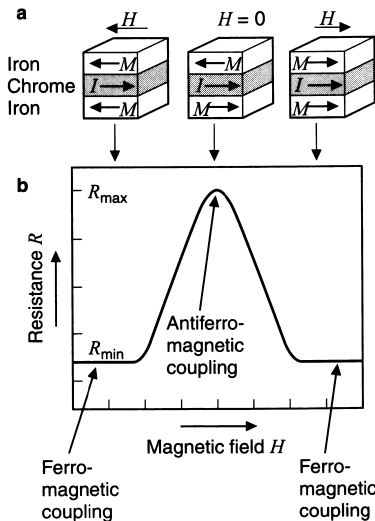
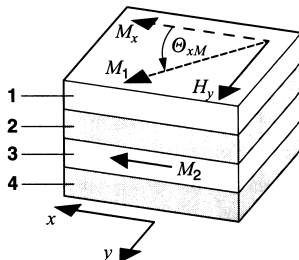


Figure 29: Design of a spin valve

- 1 Soft-magnetic layer, rotatable magnetization M_1 ,
 - 2 "Spacer", metallic non-magnetic individual layer,
 - 3 Ferromagnetic layer, fixed magnetization M_2 ,
 - 4 Antiferromagnetic coupling layer.
- M Magnetization,
 H Magnetic field,
 Θ_{xM} Angle between x and M .



sponds to the orientation of M_1 without an external magnetic field H_y . The resulting characteristic curve is shown in Figure 30; this characteristic curve shows for small magnetic fields H_y an at least piecewise linear shape – however hysteresis occurs with higher magnetic fields H_y . The maximum relative change of R for GMR spin-valve sensors is in the range of approximately 4%.

GMR sensor elements are frequently designed in a Wheatstone bridge circuit – i.e. as a full bridge.

The advantages of GMR sensors in comparison with AMR sensors lie in their higher sensitivity S ($= dR/dH$) and smaller space requirement, i.e. more signal per surface area. Furthermore, a higher resolution can be achieved on account of the smaller GMR sensor elements. All this permits the use of smaller magnets for angle sensors and smaller gear wheels for example for rpm sensors.

In addition, angle sensors with a measurement range of 360° are possible with GMR sensors, while only 180° can be achieved with AMR sensors. The reason for this is that with GMR sensors the orientation of the pinned magnetization can be arranged differently for the individual GMR sensor elements.

Tunnel magnetoresistive effect

In the case of the tunnel magnetoresistive effect (TMR effect), as with the GMR effect, the electrical resistance R changes with the orientation of the magnetization M of two

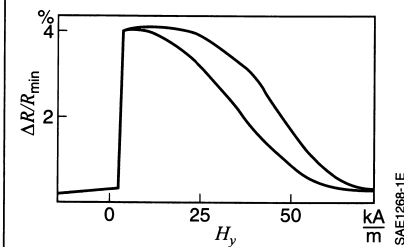
closely neighboring ferromagnetic layers. However, with TMR sensors, the current I does not, as with customary CIP-GMR sensors, flow parallel (compare here Figure 28), but instead flows due to the tunnel effect, perpendicularly to the electrically insulating and non-magnetic intermediate layer. As with the AMR and GMR effects, only applied magnetic fields in the layer plane are relevant with regard to the TMR effect.

Much smaller and even more sensitive sensors compared with GMR and AMR sensors can be designed with TMR technology; this is due to the much higher electrical resistance R for comparative relative resistance changes. The electrical power demand is reduced here in accordance with $P = U^2/R$.

References

- [1] E. Philippow: Grundlagen der Elektrotechnik. Verlag Technik, Huss-Medien, 10th Ed., 2000.
- [2] D. Zastrow: Elektrotechnik – Ein Grundlagenlehrbuch. Verlag Springer Vieweg, 19th Ed., 2014.
- [3] M. Albach: Grundlagen der Elektrotechnik 1. Pearson Studium, 3rd Ed., 2011.
- [4] M. Albach: Grundlagen der Elektrotechnik 2. Pearson Studium, 2nd Ed., 2011.
- [5] D. Meschede: Gerthsen Physik. 25th Ed., Verlag Springer Spektrum, 2015.
- [6] http://www.uni-magdeburg.de/exph/messtechnik1/Parameter_Thermoelemente.pdf.
- [7] E. Kneller: Ferromagnetismus. 1st Ed., 1962; Reprint 2012, Springer-Verlag.

Figure 30: Example of a characteristic curve of a spin-valve GMR sensor element
 H Magnetic field,
 R Resistance.



Electronics

Basic principles of semiconductor technology

Electric conductivity in solid bodies

An individual material's capacity for conducting electricity is determined by the number and mobility of the free charge carriers which it contains. The disparities in the electric conductivities displayed by various solid bodies at room temperature extend through a range defined by 10 to the 24th power. Accordingly, materials are divided into three electrical groups.

Conductors (metals)

All solid bodies contain approximately 10^{22} atoms per cubic centimeter; they are held together by electrical forces. In metals, the number of free, i.e. non-attached, charge carriers is extremely high (one to two free electrons per atom). The free carriers are characterized by moderate mobility. The electrical conductivity of metals

(e.g. silver, copper, aluminum) is high. In good conductors, this amounts to approx. 10^6 S/cm.

Non-conductors (insulators)

The number of free charge carriers in insulators (ex. aluminum oxide, Teflon, fused quartz) is practically zero. Accordingly, the electric conductivity is negligible. The conductivity of good insulators is approximately 10^{-18} S/cm.

Semiconductors

The amount of electrical conductivity in semiconductors (ex. germanium, silicon, gallium arsenide) is between that of metals and insulators. It is – unlike the conductivity of metals and insulators – greatly dependent on the following factors:

- the pressure affects the mobility of charge carriers,
- the temperature influences the number and mobility of charge carriers,
- the action of light also influences the number of charge carriers,
- the presence of additives also determines among others the number and type of charge carriers.

Because they are dependent on the above factors, semiconductors are also suitable for use as pressure, temperature and light sensors.

Doping of semiconductors

Doping, i.e. the controlled addition of electrically active foreign substances to the base material, makes it possible to define and localize the semiconductor's conductivity. This procedure forms the basis of semiconductor devices. The producible and also adjustable electric conductivity of silicon that can be reproduced by doping is 10^4 to 10^{-2} S/cm.

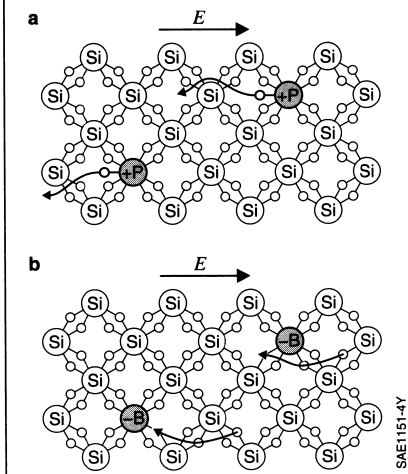
Electric conductivity of semiconductors

The following discussion focuses on the silicon-based semiconductor. In its solid state, silicon assumes the form of a crystal lattice with four equidistant contiguous atoms. Each silicon atom has four outer electrons, with two shared electrons form-

Figure 1: Doped silicon

- a) *n*-doped silicon,
b) *p*-doped silicon.
o Electron, Si Silicon, P Phosphorus, B Boron,
E Electric field.

The curved arrows indicate the direction of motion of the electrons when electric field E is applied.



ing the bond with the contiguous atoms. In this ideal state, silicon has no free charge carriers; thus it is not conductive. The situation changes dramatically with the addition of appropriate additives and the application of energy.

Doping will be explained here by reference to a clear model presentation. However, it is important to bear in mind that not all effects can be explained by reference to this model [2].

n Doping

As only four electrons are required for bonding in a silicon lattice (Figure 1a), the introduction of foreign atoms with five outer electrons (e.g. phosphorus) results in the presence of free electrons. Thus, each additional phosphorus atom will provide a free, negatively charged electron, where a single positively charged phosphorus nucleus remains. The silicon is transformed into an *n* conductor (*n*-type silicon), since there is an excess of negative charges (electrons). In response to an externally applied voltage, an electric field E is generated, which gives the mobile charge carriers a preferred direction for movement (Figure 1).

p Doping

The introduction of foreign atoms with three outer electrons (e.g. boron) produces electron holes which result from the fact that the boron atom has one electron less for complete bonding in the silicon lattice (Figure 1b). This hole indicates a missing electron. Holes remain in motion within the silicon; they are filled by electrons that leave behind their own hole. In an electric field, they travel in the opposite direction of the electrons. The holes exhibit the properties of a free positive charge carrier. Thus, every additional boron atom provides a free, positively charged electron hole (positive hole). The silicon is transformed into a *p* conductor and is called a *p*-type silicon.

Intrinsic conduction

Heat and light also generate free mobile charge carriers in undoped silicon; the resulting electron-hole pairs produce intrinsic conductivity in the semiconductor material. This conductivity is generally

modest in comparison with that achieved by doping. Increases in temperature induce an exponential rise in the number of electron-hole pairs, ultimately obviating the electrical differences between the *p* and *n* regions produced by doping. This phenomenon defines the maximum operating temperature to which semiconductor devices may be subjected: This is for germanium 90 to 100 °C, for silicon 150 to 200 °C, and for gallium arsenide 300 to 350 °C.

A small number of opposite-polarity charge carriers is always present in both *n*-type and *p*-type semiconductors. These minority charges exert a considerable influence on the operating characteristics of virtually all semiconductor devices.

pn junction in the semiconductor

The area of transition between a *p*-type zone and an *n*-type zone within the same semiconductor crystal is referred to as the *pn* junction. The properties of this area exercise a major influence on the operating properties of most semiconductor devices.

pn junction without external voltage

The *p*-type zone has numerous holes, and the *n*-type zone has only very few; there is a very large number of electrons in the *n*-type zone, while in the *p*-type zone there is an extremely small number. Each type of mobile charge carrier tends to move across the concentration gradient, diffusing into the other zone (Figure 2b).

The diffusion of holes into the *n*-type zone gives the *p*-type zone in the *pn* junction area a negative charge, since the negatively charged core atoms (i.e. boron atoms) remain stationary. Electron depletion results in the *n*-type region being positively charged, since here this is an excess of stationary, positively charged atomic residues (e.g. phosphorus). The result is an electrical potential between the *p*- and *n*-type zones (diffusion potential U_D). This potential opposes the respective migration tendencies of the charge carriers. This ultimately brings the exchange of holes and electrons to a halt. The potential current U_D created on account of diffusion cannot be directly measured from the outside and in silicon is typically just 0.6 V.

A poorly conducting zone deficient in mobile charge carriers is thus produced at the *pn* junction: This zone is called the space-charge region or depletion layer. This zone has an electric field whose strength is also dependent on the externally applied voltage.

pn junction with external voltage

Now the conditions at a diode will be explained because a *pn* junction corresponds to the structure of a diode; here the anode is situated at the *p*-doped silicon, while the cathode is situated at the *n*-doped zone.

The application of voltage *U* in the reverse direction (negative pole at the *p*-type zone and positive pole at the *n*-type zone) extends the space-charge region (Figure 2c). Consequently, the flow of current *I* is blocked except for a minimal residual current (reverse current) which stems from the minority charge carriers. The voltage *U* then drops within the space-charge region; this region is therefore subject to a high electric field strength.

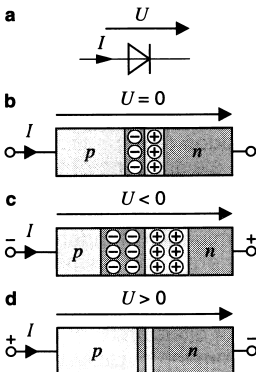
Breakdown voltage is the level of reverse-direction voltage beyond which a minimal increase in voltage will suffice to

produce a sharp rise in reverse current (Figure 3). This effect can be explained as follows: Electrons which reach the space-charge region are greatly accelerated on account of the high field strength. In this way, they can for their part generate free charge carriers as a result of impacts; this is also known as collision ionization. This brings about a dramatic rise in current and results in avalanche breakdown. In addition to avalanche breakdown, there is also Zener breakdown, which is based on the tunnel effect. The breakdown of a *pn* junction can destroy it and for this reason is often not wanted. In many cases, however, breakdown is wanted. Avalanche breakdown and Zener breakdown only occur when the diode is operated in the reverse direction.

The application of a voltage *U* in the forward direction (positive pole at the *p*-type zone and negative pole at the *n*-type zone) reduces the space-charge region (Figure 2d). Charge carriers permeate the *pn* junction, resulting in a large flow of current in the forward direction (Figure 3), as the space-charge region no longer has a significant resistance. Only the bulk resistance, i.e. the ohmic resistance of the doped layers, is effective. The current *I* increases exponentially as a function of *U*. It is, however, important to be aware of "thermal breakdown", at which point the semiconductor can be destroyed on account of the high level of heat. This can occur, for example, if the diode is operated in the forward direction with an unacceptably high current.

Figure 2: *pn* junction in a diode

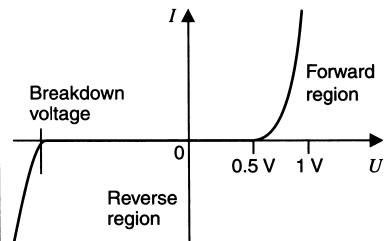
- a) Diode circuit symbol,
 - b) *pn* junction without external voltage,
 - c) *pn* junction in reverse direction,
 - d) *pn* junction in forward direction.
- U* Applied voltage (diode voltage),
I Diode current.
⊕ Positively charged atomic residues,
⊖ Negatively charged atomic residues.



SAE1153-2Y

Figure 3: Characteristic curve of a silicon diode

- U* Applied voltage (diode voltage),
- I* Diode current.



SAE1154-3E

Discrete semiconductor devices

The properties of the *pn* junction and the combination of several *pn* junctions in a single semiconductor-crystal chip provide the basis for a steadily increasing array of inexpensive, reliable, rugged, compact semiconductor devices. A single *pn* junction forms a diode, two *pn* junctions are used for transistors. The planar technique makes it possible to combine numerous operating elements on a single chip to form the important component group known as integrated semiconductor circuits. Semiconductor chips measure no more than several square millimeters and are usually installed in standardized housings (metal, ceramic or plastic).

Diodes

A diode is a semiconductor device incorporating a single *pn* junction. Its specific properties are determined by the distribution pattern of the dopant in the crystal. Diodes which conduct currents in excess of 1 A in the forward direction are frequently referred to as power diodes.

Rectifier diode

The rectifier diode acts as a form of current valve; it is, therefore, ideally suited to rectifying alternating current. The forward current can be approximately 10^7 times higher than the current in the reverse direction (reverse current, Figure 3). As the temperature rises, the power of this current increases significantly.

Diodes can also be used as polarity-reversal protection diodes in order to prevent a current flow in the event of an incorrect voltage connection. It is also customary to use them as free-wheeling diodes.

Rectifier diode for high reverse voltage

In the case of a rectifier, the voltage drops over the space-charge region. Because this region is generally only a few micrometers in size, the high reverse currents produce a high electric field strength and the free electrons can accelerate greatly. Accelerated electrons can cause the semiconductor to be destroyed (avalanche breakdown). To prevent this, it has

proved useful to integrate an intrinsic layer between the *p*- and *n*-layers because this layer only contains few free electrons and thereby reduces the danger of a breakdown.

Switching diode

A switching diode is generally employed for rapid switching between high and low impedances. More rapid switching response can be achieved by diffusing gold into the material; this promotes the recombination of electrons and holes. During switching the space-charge region must be recharged accordingly; during switching-on from the blocking to the conducting state the initially charge-free space-charge region is filled with charges, which results in the forward recovery time. Accordingly, when the diode passes into the blocking state the (excess) charge carriers located in the space-charge region must be removed – the time needed for this is called the backward recovery time. For bipolar transistors the saturation ($U_{CE} < 0.2$ V) is therefore critical when rapid switching is required. The storage charge or the maximum reverse current is often also specified here.

A distinction is made in the losses between forward power losses (forward voltage times current) and switching power losses; at lower frequencies up to a few 100 Hz the forward power losses dominate while at higher frequencies the switching power losses increasingly play a role.

Power diodes

When compared with switching diodes, power diodes are designed for higher currents (> 1 A) and high reverse voltages (up to over 1 kV). These demands arise from the conditions in power electronics where high current intensities and high (reverse) voltages can occur with low forward power losses being desired. Power diodes therefore frequently have – similarly to rectifier diodes for high reverse voltages – a lightly doped “middle layer” between the *p*- and *n*-doped regions – i.e. a pin structure: *p*-doped, undoped or intrinsic, *n*-doped. The space-charge region can further expand in the lightly doped region and thus a higher (reverse) voltage

can be achieved without the dopings having to be reduced in the p and n regions.

Aside from pure silicon diodes, Schottky diodes based on silicon carbide (SiC) are also used here; these diodes exhibit a forward voltage of 0.8 V and a maximum reverse voltage of barely 2 kV. Furthermore, SiC diodes can also be operated at higher temperatures (up to 200 °C).

Free-wheeling diodes

High voltages can occur during the rapid switching of currents by inductance since the induced voltage is proportional to dI/dt . Frequently free-wheeling, or flyback, diodes are used to prevent damage to components such as power transistors (e.g. MOSFET or IGBT). In this case the current flow is not abruptly interrupted, but can continue to flow via the free-wheeling diodes so that high induced voltages are avoided.

Zener diode

A Zener diode is a semiconductor diode which, once a specific initial level of reverse voltage is reached, responds to further increases of reverse voltage with a sharp rise in current flow. This phenomenon is a result of a Zener and/or avalanche breakdown. Zener diodes are designed for continuous operation in this breakdown range. These are frequently used to provide a constant voltage or reference voltage.

Variable-capacitance diode (varactor)

The space-charge region at the pn junction functions as a capacitor; the dielectric element is represented by the semiconductor material in which no charge carriers are present. Increasing the applied voltage extends the depletion layer and reduces the capacitance, while reducing the voltage increases the capacitance.

Schottky barrier diode (Schottky diode)

A Schottky diode contains a metal-to-semiconductor junction. As the electrons move more freely from the n -type silicon into the metal layer than in the opposite direction, an electron-depleted layer is created in the semiconductor material; this is the Schottky barrier layer. Charges are carried exclusively by the electrons, a factor

which results in extremely rapid switching, as the minority carriers do not perform any charge storage function (charge-carrier surplus in the space-charge region is reduced). The forward voltage and thus the voltage drop is at approximately 0.3 V smaller in Schottky diodes than in silicon diodes (approximately 0.6 V). Because of the rapid switching and the comparably low losses, Schottky diodes are used in high-frequency circuits down into the microwave range.

Solar cell

The photovoltaic effect is applied to convert light energy directly into electrical energy. Solar cells, consisting largely of semiconductor materials, are the basic elements of photovoltaic technology. Exposure to light can cause the formation of free charge carriers (electron-hole pairs) in the semiconductor. If the semiconductor incorporates a pn junction, the free charge carriers separate in its electric field before proceeding to the metal contacts on the semiconductor's surface. Depending on the semiconductor material used, a DC voltage (photovoltage) ranging from 0.5 to 1.2 V is generated between the contacts. This occurs only when the light quanta have at least the energy required to create an electron-hole pair; the energy is proportional to the frequency and inversely proportional to the wavelength of the light. The theoretical efficiency of crystalline silicon solar cells is approximately 30 %.

Photodiode

The photovoltaic effect is utilized in a photodiode. The pn junction is operated in the reverse direction. Incident light creates additional free electrons and holes. These increase the reverse current (photovoltaic current) in direct proportion to the intensity of the light. The photodiode is thus, in principle, very similar to the solar cell.

Light-emitting diode

A light-emitting diode or LED is an electroluminescent lamp which consists of a semiconductor element with pn junction. The charge carriers (free electrons and holes) recombine during operation in forward direction. The amount of energy

released in this process is converted into electromagnetic radiation energy.

Depending on the choice of semiconductor and its doping, the LED emits in a limited spectral range. Frequently used semiconductor materials are: gallium arsenide (infrared), gallium arsenide phosphide (red to yellow), gallium phosphide (green), and indium gallium nitride (blue). In order to generate white light, either a combination of three LEDs with the primary colors Red, Green and Blue is used, or a fluorescent dye is excited with a blue or ultraviolet-emitting LED. The maximum luminous efficiency of present-day LEDs is over 300 lm/W – the maximum possible value is approximately 350 lm/W, where only the spectrum that can be detected by the human eye is taken into account here (photometric quantity “luminous flux” in lumens or abbreviated as “lm”).

OLED

OLED stands for “Organic Light-Emitting Diode”. An OLED is a very thin, two-dimensional light source which emits soft and for the most part non-dazzling light – sharp cast shadows do not occur. OLEDs are in a good approximation Lambertian emitters, i.e. they always appear with uniform brightness regardless of the viewing angle. The small geometric dimensions are also to be noted; thus, OLEDs are less than 2 mm thick – including glass substrate and encapsulation. A further particularity is the flexibility of the extremely flat design, which creates further possible applications in addition to the light properties with high color rendering.

OLEDs - as non-dazzling large-area emitters with a broad spectrum - are particularly suitable for interior applications in automobiles. Constant-current operation suggests itself as the form of activation, enabling light changes on account of aging, temperature or even by production tolerances to be minimized. OLEDs can be dimmed for example by pulse-width modulation.

During the manufacture of OLEDs a number of thin organic semiconductor layers are placed in succession on the conductive substrate (formation of a stack); these layers can form polymeric

substances or “small-molecule materials” as an amorphous layer.

Further conductive and transparent electrodes, which cover the organic layers on both sides, are then used. Oxides such as indium tin oxide (ITO) are used as the material for these electrodes; on account of the non-ideal conductivity there is a voltage drop in the active layers from the outside inwards, which results in a corresponding reduction of the radiation intensity from the outside inwards. A remedy can be provided by an auxiliary structure made of more conductive material. However, the homogeneity of the radiation intensity can be improved by the multi-layer design (stacking).

The maximum luminous efficiency for OLEDs is today approximately 100 lm/W – i.e. still significantly below that of LEDs. However, the service life must also be noted; this decreases with higher operating current – and consequently higher light output.

Laser LED

New lighting systems increasingly use laser LEDs – semiconductor-based laser light sources. These are small spotlights with very high luminance (a factor of approximately 4 compared with LEDs). It is to be noted here that the initially generated blue laser beam ($\lambda \approx 450$ nm) does not leave the headlamp, but is converted with the aid of luminescent materials (ceramic high-power materials) into visible white light with a color temperature of around 5,500 K. A light cone with a very high range (up to 600 m) can then be achieved with very small optical lenses. This makes possible small headlamps with high light output. The performance of a laser LED is decisively influenced by the properties of these luminescent materials and their optical properties. Hybrid light concepts are frequently used at present so that LEDs and laser LEDs are used for the driving lights and for the high-beam lights respectively.

Bipolar transistors

Two contiguous *pn* junctions produce the transistor effect, a feature employed in the design of components used to amplify electrical signals and to assume switching duties. Bipolar transistors consist of three zones of varying conductivity, the configuration being either *pn*p or *n*p*n*. The zones (and their terminals) are called emitter E, base B, and collector C (Figure 4).

There are different transistor classifications, depending on the fields of application: small-signal transistors (power dissipation up to 1 watt), power transistors, switching transistors, low-frequency transistors, high-frequency transistors, microwave transistors, and phototransistors. They are termed bipolar because charge carriers of both polarities (holes and electrons) are involved in the transistor effect.

Operation of a bipolar transistor

The operation of a bipolar transistor is explained here using an *n*p*n* transistor as the example (Figure 5). The *pn*p transistor is obtained similarly by switching the *n*- and *p*-doped zones. The base-emitter junction is forward-biased; this is shown in Figure 4b as a diode between base B and emitter E. In this way, with sufficient voltage U_{BE} , electrons are injected into the base zone and base current flows.

The base-collector junction is reverse-biased; this is shown in Figure 4b as a diode between base B and collector C. This creates a space-charge region at the *pn* junction between base and collector with a high electric field.

Because of the reverse-biased diode between the base and the emitter, a high current consisting of electrons flows from the emitter to the base. Here, however, only a small fraction can recombine with the (much fewer) holes and flow as base current I_B out of the base terminal; what must be borne in mind is that the technical current direction – i.e. the direction of motion of the positive charge carriers – is specified in Figure 4. The much greater amount of electrons injected into the base diffuses through the base zone to the base-collector junction and then flows as collector current I_C to the collector (Figure 5). Because the base-collector diode

Figure 4: *n*p*n* transistor

- a) Diagram,
b) Structure.
E Emitter,
B Base,
C Collector.

U_{BE} Base-emitter voltage,
 U_{CE} Collector-emitter voltage.
 I_B Base current,
 I_C Collector current,
 I_E Emitter current.

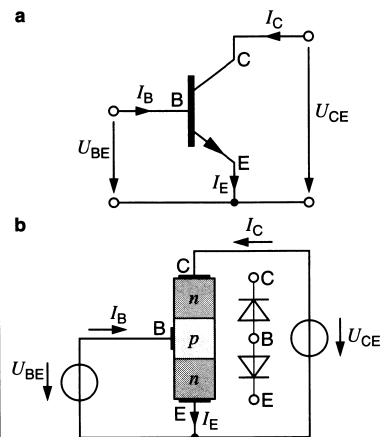
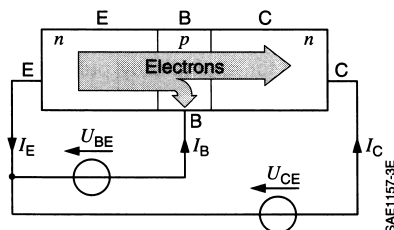


Figure 5: Operation of an *n*p*n* transistor

- E Emitter,
B Base,
C Collector.

U_{BE} Base-emitter voltage,
 U_{CE} Collector-emitter voltage,
 I_B Base current,
 I_C Collector current,
 I_E Emitter current.



is operated in the reverse direction and a space-charge region is predominant, virtually all (approximately 99 %) of the electrons flowing from the emitter are "drawn off" through the strong electric field in the space-charge region from the collector. In this case, a linear connection applies approximately between the collector current I_C and the base current I_B :

$$I_C = B I_B$$

where B is the current gain, which is generally between 100 and 800. In the bipolar transistor, the following relation for the emitter current I_E also applies (cf. Figure 4 and Figure 5):

$$I_E = I_B + I_C.$$

Assuming that I_B on account of the current gain B is much smaller than I_C , it then follows:

$$I_E \approx I_C.$$

The very thin (and relatively low-doped) base represents a barrier that can be adjusted by means of the base-emitter voltage U_{BE} from the charge-carrier flow from the emitter to the collector. With a small change of U_{BE} and the base current I_B , it is possible to control a greater change

of the collector current I_C and the collector-emitter voltage U_{CE} . Small changes in the base current I_B thus bring about large changes in the emitter-collector current I_C . The *npn* transistor is a bipolar, current-controlled semiconductor amplifier. In all, there is a power amplification.

The output characteristic curve for an *npn* transistor is shown in Figure 6. From the saturation voltage of approximately 0.2 V for U_{CE} , the collector current I_C is virtually only dependent on the base current I_B as the parameter; this region is termed the "active region": U_{CE} here has virtually no influence on I_C and the following applies:

$$I_C = B I_B.$$

The region below the saturation voltage is called the "saturation region". In this region, I_C increases sharply with U_{CE} .

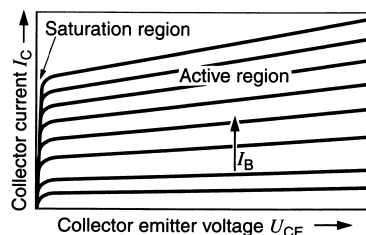
Darlington transistor

The Darlington arrangement of two bipolar transistors is often used to achieve a higher current amplification; the interconnection of the two transistors is also called a Darlington transistor (Figure 7). Transistor T_1 (as emitter follower) switches transistor T_2 so that in total a very high current amplification β (up to 50,000) is achieved as the product of the two current amplifications of the two transistors T_1 and T_2 : $\beta = \beta_1 \beta_2 (\approx I_{C2}/I_{B1})$. Compared with power bipolar transistors ($\beta \approx 5 \dots 10$), the current amplification is much higher and consequently the control current is much lower.

Darlington transistors are used when a voltage (which is not to be subjected to load) connects as high a load as possible (i.e. current). On account of the phase displacement between the input (control current I_{B1} to the base of T_1) and the output (collector current I_{C2} in T_2), the Darlington transistor is only suitable for relatively low frequencies – this circuit is therefore not used in high-frequency circuits. The main reason for this is the time required for the switching of T_2 through the removal of the excess charges in the base of T_2 via resistance R ; if R is selected as too small, the total current amplification β is reduced.

Figure 6: Output characteristic curve of an *npn* transistor

U_{CE} Collector-emitter voltage,
 I_C Collector current,
 I_B Base current as parameter
of characteristic curve.



SAE1160-1E

Another drawback is the voltage at the input of 1.2...1.4 V, which is twice as high compared with simple bipolar transistors. The forward voltage U_{CE2} at T_2 increases by U_{BE2} to approximately 0.9 V (compared with the saturation voltage of 0.2 V for simple bipolar transistors – here U_{CE1}) or approximately 2 V for power types. As a result the losses are increased (proportionally $U_{CE2} \cdot I_{C2}$).

Diode D_3 is a free-wheeling diode for inductive loads in order to avoid high voltages between the emitter and collector of T_2 .

IGBT

The insulated-gate bipolar transistor (IGBT) is an important switching element in the middle power range for currents up to 3 kA and voltages up to 3 kV. The IGBT is characterized by relatively low switching power losses and high efficiency.

The IGBT combines the advantages of a MOSFET (T_1 in Figure 8, see Field-effect transistors) with low-power, voltage-controlled activation and of a power bipolar transistor T_2 with a relatively low forward voltage at the output: In Figure 8 voltage U_{GE} is used at the MOS transistor T_1 to switch on the *pnp* bipolar transistor T_2 , which then directs the output current I_C with relatively low forward voltage U_{CE} ; the saturation voltage U_{CESAT} is in the range of 2...3 V.

Due to the curved decay of the current during switch-off on account of the light doping, the switching frequency is limited – as a rule 20 kHz to max. 100 kHz for smaller power levels.

Field-effect transistors

In a field-effect transistor (FET), control of the current flow in a conductive path is essentially exercised by an electric field. The field, in turn, is generated by a voltage applied at a gate (Figure 9). Field-effect transistors differ from their bipolar counterparts in that they utilize only a single type of charge carrier (either electrons or holes), giving rise to the alternate designation “unipolar transistors”. These are subdivided into the following categorizations: junction-gate field-effect transistors (junction FET, JFET) and insulated-gate field-effect transistors, particularly MOS field-effect transistors (MOSFET or MOS transistors).

MOS field-effect transistors are well suited for application in high-integration circuits. Power field-effect transistors represent a genuine alternative to bipolar power transistors in many applications.

The advantages of a bipolar transistor and of a field-effect transistor are utilized in power electronics in “insulated-gate bipolar transistors” (IGBT), which exhibit a low volume resistance (for small losses) and comparatively low triggering power.

Figure 7: Circuit of a Darlington transistor

T Bipolar transistors, R Resistor, E Emitter, B Base, C Collector, D Diode.
 I_C Collector current,
 U_{BE} Base-emitter voltage,
 U_{CE} Collector-emitter voltage.

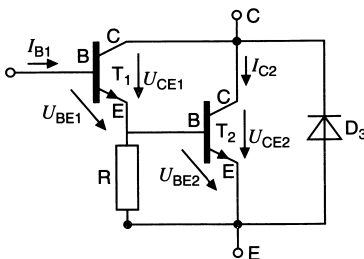
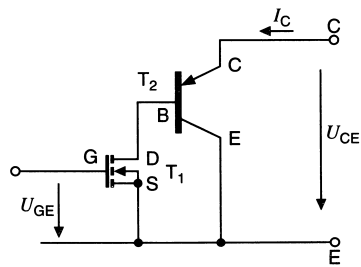


Figure 8: Circuit of an IGBT

T_1 Field-effect transistor,
 T_2 Bipolar transistor.
G Gate,
S Source,
D Drain.
 U_{GE} Gate-emitter voltage,
 I_C Collector current.



Operation of a junction FET

The operation of a junction-gate field-effect transistor is explained by reference to the n channel type (Figure 9). The terminals of a field-effect transistor are referred to as gate (G), source (S) and drain (D).

Positive direct voltage U_{DS} is applied at the ends of an n -type crystal. Electrons flow through the channel from the source to the drain. The width of the channel is defined by two laterally diffused p -type zones and by the negative gate-source voltage U_{GS} applied at them. The voltage U_{GS} between the control electrode (gate G) and the source terminal thus controls the current I_D between the source and the drain.

Only charge carriers of one polarity are required for field-effect transistor operation. The power necessary for controlling the current is virtually nil. Thus, the junction FET is a unipolar, voltage-controlled component. Increasing U_{GS} causes the space-charge regions to extend further into the channels, thereby constricting the channel and thus the current path (see dashed lines in Figure 9). If the voltage U_{GS} at the gate is zero, the channel between the two p -type zones is not constricted and the current I_D from the drain to the source is at its maximum.

Figure 9: Junction-gate field-effect transistor with n channel

- a) Diagram,
b) Structure.

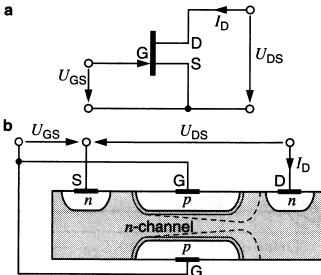
The lightly shaded region around the source and drain contacts is more heavily doped than the channel.

G Gate, S Source, D Drain.

U_{DS} Drain-source voltage,

U_{GS} Gate-source voltage,

I_D Drain current.



SAE1159-2E

The transfer characteristic curve – i.e. I_D as a function of U_{GS} – then looks exactly the same as the characteristic curve of a self-conducting n -channel MOSFET as shown in Figure 11c.

Operation of an MOS transistor

The operation of an MOS transistor (metal-oxide semiconductor) is explained by reference to the self-blocking n -channel MOSFET (enhancement type) (Figure 10). If no voltage is applied to the gate, then no current will flow between the source and the drain: the pn junctions remain in blocking mode. The application of a positive voltage at the gate causes, on account of the electrostatic induction in the p -type zone below this gate, the holes to be displaced toward the interior of the crystal and electrons – which are always present in p -type silicon as minority charge carriers – to be pulled to the surface. A narrow n -type layer, an n channel, forms below the surface. Current can now flow between the two n -type zones (source and drain). This current consists exclusively of electrons. As the gate voltage acts through an insulating oxide layer, no stationary current flows through the gate; no power is required for the control function. It is required merely to activate and deactivate electrical power in order to recharge the gate capacity. In summary, the MOS transistor is a unipolar, voltage-controlled component.

In the case of the self-conducting n -channel MOSFET (depletion type, Figure 11a), the gate-source voltage U_{GS}

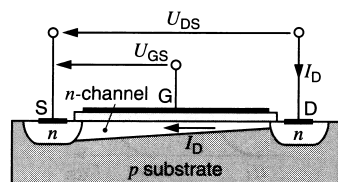
Figure 10: n -channel MOSFET, cross-section

S Source, G Gate, D Drain.

U_{DS} Drain-source voltage,

U_{GS} Gate-source voltage,

I_D Drain current.



SAE1161-1E

is between the here negative threshold voltage U_T and zero volts (Figure 11c). At $U_{GS} = 0\text{ V}$, the self-conducting n -channel MOSFET has a channel below the gates for the current flow. Figure 11c shows I_D as a function of U_{GS} , where the circuit is operated in the active region with sufficient and constant U_{DS} . The transfer characteristic curve is a parabola. In contrast, the self-blocking n -channel MOSFET (Figure 11b) conducts only from the positive threshold voltage $U_T^* > 0\text{ V}$ here (see Figure 11c). The self-blocking MOSFET is much more common than the self-conducting MOSFET.

The output characteristic curve of a self-blocking n -channel MOSFET is shown in Figure 12. The region below the knee-point voltage U_K , i.e. for $U_{DS} < U_K$ is referred to, on account of the linear characteristic curve, as the linear or ohmic region; here the MOSFET behaves like an ohmic resistance. Above the knee-point voltage U_K , i.e. for $U_{DS} > U_K$, the output current I_D is virtually uninfluenced by the drain-source voltage U_{DS} ; this region is

known as the cut-off region. The amount of I_D is dependent only on the gate-source voltage U_{GS} . The formula is:

$$I_D = 0.5\,K\,(U_{GS} - U_T)^2$$

where K is the proportionality number (dependent among others on technological quantities) and U_T the threshold voltage from which the transistor conducts, i.e. a channel forms (see Figure 11c).

PMOS, NMOS, CMOS transistors

As well as the n -channel MOSFET (NMOS transistor), mixing the doping produces the PMOS transistor. As the electrons in the NMOS transistor are more mobile, it operates more rapidly than the PMOS device, although the latter was the first to become available due to the fact that it is physically easier to manufacture.

It is also possible to employ complementary MOS technology to pair PMOS and NMOS transistors in a single silicon chip; the resulting devices are called complementary MOS transistors (CMOS transistors, Figure 13). The specific advantages of the CMOS transistor are extremely low power dissipation, a high degree of immunity to interference, relative insensitivity to varying supply voltages, and suitability for analog signal processing and high-integration applications [2].

Figure 11: n -channel MOSFET

- a) Diagram of self-conducting n -channel MOSFET,
 - b) Diagram of self-blocking n -channel MOSFET,
 - c) Characteristic curves.
1 Characteristic curve of self-conducting n -channel MOSFET,
2 Characteristic curve of self-blocking n -channel MOSFET.
- U_{GS} Gate-source voltage,
 I_D Drain current,
 U_{DS} Drain-source voltage,
 U_T, U_T^* Threshold voltage.

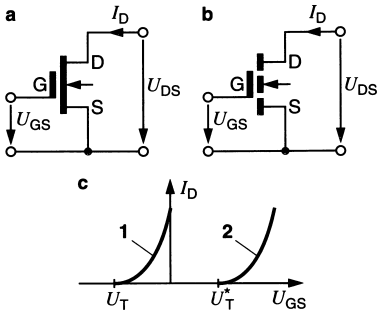
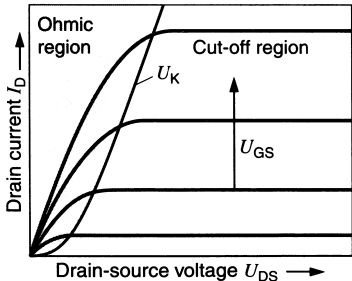


Figure 12: Output characteristic curve of a self-blocking n -channel MOSFET

- U_{DS} Drain-source voltage,
 U_{GS} Gate-source voltage,
 I_D Drain current,
 U_K Knee-point voltage.



BCD hybrid technology

Integrated structures for power-electronics applications are becoming increasingly important. Such structures are realized by combining bipolar and MOS components on a single silicon chip, thereby utilizing the advantages of both technologies. BCD hybrid technology is a significant manufacturing process in automotive electronics and also facilitates the manufacture of MOS power components (DMOS). This technology is a combination of bipolar, CMOS and DMOS technologies [2].

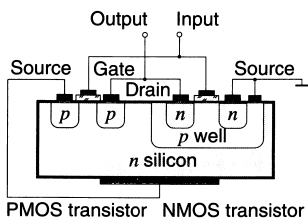
Operational amplifiers

Areas of application of operational amplifiers

The name "operational amplifier" (OPA) comes from analog computing technology and denotes an (almost) ideal amplifier. Because of its properties, it was used particularly in analog computers to solve nonlinear differential equations, i.e. as a summator, integrator and differentiator. The rapid development of digital electronics saw analog computers being increasingly driven from the market with the result that today analog computers do not play a role at all.

By integrating them in microelectronic circuits, it is today possible to offer such operational amplifiers at a very economical price on the market so that many amplifier applications can be put into effect. To achieve the desired properties, operational amplifiers in their integrated form contain some (depending on the requirement 10 to 250) transistors, the number of which however is only of subordinate importance with regard to integration.

Figure 13: CMOS inverter composed from PMOS and NMOS technology



UAE0461-2E

Starting out from a "normal" operational amplifier with voltage input and voltage output (VV operational amplifier), first the behavior of an ideal operational amplifier will be described and its use demonstrated. Then the real – i.e. actual – properties will be examined in more detail and its effect to be realized on the circuit analyzed.

Basic principles

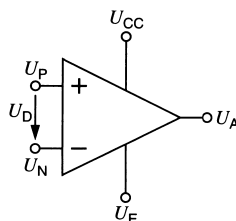
The ideal standard operational amplifier is an amplifier with two inputs and (normally) one output (Figure 14). The inputs are the non-inverting input and the inverting input. The differential voltage U_D is amplified and then made available at the output as the output voltage U_A . The following formula applies:

$$U_A = A_D U_D.$$

A_D is the open-loop gain. The operational amplifier is connected to a positive supply voltage and to a negative supply voltage with regard to the frame potential. In the case of a unipolar supply, the negative supply voltage can be applied to the frame potential. The supply voltages are not usually indicated in many circuit diagrams. They are however very much needed in order to guarantee the energy supply to the operational amplifier.

Figure 14: Basic diagram of an operational amplifier

+ Non-inverting amplifier input,
 – Inverting amplifier input,
 U_D Differential voltage between the two input potentials U_P and U_N where
 $U_D = U_P - U_N$,
 U_A Output voltage,
 U_{CC} positive supply voltage,
 U_E negative supply voltage.



SAE1164Y

The following variants of the operational amplifier are common (Figure 15):

- “Normal” operational amplifier (VV operational amplifier) with voltage inputs and voltage output.
- Transconductance amplifier (VC operational amplifier) with voltage inputs and current output.
- Transimpedance amplifier (CV operational amplifier) with current inputs and voltage output.
- Current amplifier (CC operational amplifier) with current inputs and current output.

Figure 15: Operational-amplifier types (diagram)

- Normal operational amplifier (VV) with voltage input and voltage output.
- Transconductance amplifier (VC) with voltage input and current output.
- Transimpedance amplifier (CV) with current input and voltage output.
- Current amplifier (CC) with current input and current output.

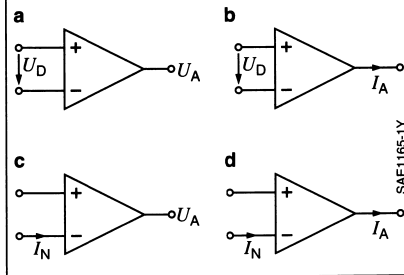
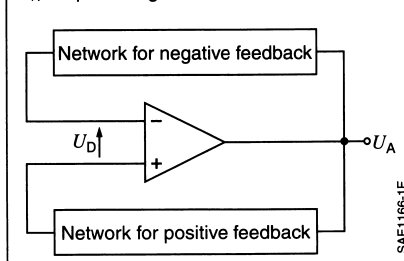


Figure 16: Negative and positive feedback

U_D Differential voltage,
 U_A Output voltage.



As a rule, the VV operational amplifier is used; this will now be explained in more detail in the following. Because the circuit arrangement is of crucial importance to the function of an operational amplifier, this will be addressed in more detail first. The distinction between positive feedback and negative feedback is important here. Furthermore, an ideal operational amplifier will be assumed when deriving the correlations.

Circuit arrangement: negative and positive feedback

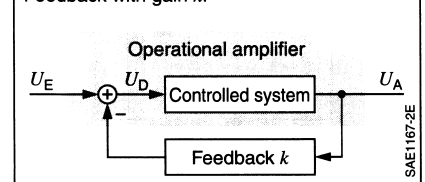
Negative feedback counteracts the cause. In the operational amplifier, a connection is required from the output to the inverting input for this purpose (Figure 16). This connection can be implemented by a network. The cause of a change in the output voltage U_A is always a change in the differential voltage U_D at the input; negative feedback therefore always acts in such a way that the voltage U_D becomes small and ideally zero.

Positive feedback, in contrast to negative feedback, supports the cause of the change at the output. U_A is thus amplified by positive feedback, i.e. U_D increases as U_A changes and is therefore always not equal to zero. In this way, the output voltage U_A can assume only two stationary values, namely the maximum value or the minimum value.

From a control-engineering standpoint, a negative feedback results from the operational amplifier and the feedback as shown in Figure 17. Taking a high gain A_D into consideration, it follows:

Figure 17: Negative feedback

U_E Input voltage as setpoint value,
 U_A Output voltage,
 U_D Differential voltage is obtained at the summing point $\oplus$ $U_D = U_E - k U_A$.
Controlled system with gain A_D ,
Feedback with gain k .



$$U_A = A_D U_D = A_D (U_E - k U_A)$$

and for the total gain

$$A = \frac{U_A}{U_E} = \frac{A_D}{1 + k A_D} \approx \frac{1}{k}.$$

It thus becomes clear that, in spite of a very high open-loop gain A_D on the part of the operational amplifier, with the aid of negative feedback a finite gain A can be set with the negative-feedback network. This is explained below in greater detail using examples.

Ideal and real operational amplifier

First, the properties of an ideal operational amplifier as shown in Figure 18 will be summarized. For further information, see [1].

- Common-mode input resistance between, in each case, one input and ground, where: $r_{GL,P} = U_P/I_P$; $r_{GL,N} = U_N/I_N$. Generally speaking, r_{GL} can be ignored.
- Differential input resistance between the two inputs; here: $r_D = (U_P - U_N)/I_P$. r_D is increased by negative feedback.
- Output resistance, differential quantity $r_A = dU_A/dI_A$. r_A is decreased by negative feedback.
- Offset voltage U_{OS} : Characteristic quantity for describing the fact that even in the event of a short circuit between the two inputs (i.e. $U_D = 0$) the output voltage U_A is not equal to zero.
- Common-mode rejection ratio (CMRR): This quantity describes the change in the output voltage U_A if the two input voltages U_P and U_N change simultaneously (in the case of periodic input signals co-phasally), i.e. U_D remains constant.
- Power-supply rejection ratio (PSRR): Change in the output voltage U_A on account of a change in the supply voltages.

The essential idealizations are:

- The open-loop gain A_D approaches infinity; in the case of negative feedback the following applies: $U_D = 0$.
- The input currents I_N and I_P each approach zero.
- If I_N and I_P each approach zero, it follows that the common-mode and dif-

ferential input resistances approach infinity.

- The offset voltage U_{OS} approaches zero.
- The output resistance R_A approaches zero.
- The common-mode rejection ratio (CMRR) approaches infinity, i.e. in the case of an equal and co-phasal change in the voltages U_P and U_N , U_A remains unchanged.
- The power-supply rejection ratio (PSRR) approaches infinity, i.e. in the case of a change in the supply voltage, U_A does not change.
- The behavior is not dependent on the frequency.

In reality, the above-mentioned idealizations do not apply entirely:

- The open-loop gain A_D is in the range of 10^4 to 10^7 .
- The input currents I_N and I_P are in the range of 10 pA to 2 μ A.
- The common-mode input resistance is in the range of 10^6 to $10^{12} \Omega$, the differential input resistance at up to $10^{12} \Omega$.
- The output resistance R_A is in the range of 50 Ω to 2 k Ω .
- The common-mode rejection ratio (CMRR) is in the range of 60 to 140 dB.
- The power-supply rejection ratio (PSRR) is in the range of 60 to 100 dB.
- The behavior is dependent on the frequency (low-pass behavior).

Figure 18: Ideal operational amplifier

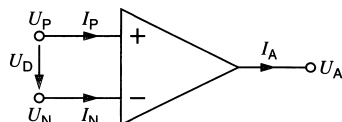
U_D Differential voltage between the two input potentials U_P and U_N where

$$U_D = U_P - U_N,$$

I_P, I_N Input currents,

U_A Output voltage,

I_A Output current.



Basic circuits

The external circuit arrangement of an operational amplifier determines the behavior of the entire circuit. Here, negative feedback plays the dominant role, since it enables the gain to be set exactly through the choice of resistances. The function shall now be explained by reference to several examples.

Inverting amplifier

The basic circuit for an inverting amplifier is shown in Figure 19. The name can be attributed to the negative gain, i.e. in the case of a periodic input voltage, the output voltage U_A is always in phase opposition to the input voltage U_1 . In the following, it is important that, on account of the negative feedback and the high open-loop gain A_D , the differential voltage U_D at the input be constantly zero, since the non-inverting and inverting inputs are kept at the same potential. Since, on account of the negative feedback, the differential voltage U_D is regulated to zero, this is referred to as a "virtual short circuit". "Virtual ground" is also referred to here, as the inverting input is actively kept at zero (i.e. at frame potential). In addition, the input currents are ignored, in particular $I_N = 0$ is set. The following applies:

$$I_{R1} = \frac{U_1}{R_1} \text{ and } I_{R2} = -\frac{U_A}{R_2}.$$

Where $I_{R1} = I_{R2}$ it then follows:

$$U_A = -\frac{R_2}{R_1} U_1.$$

The output voltage U_A is thus directly dependent on the input voltage U_1 and on the choice of resistances R_2 and R_1 .

Non-inverting amplifier

The non-inverting amplifier can be handled along similar lines to the inverting amplifier (Figure 20). Due to negative feedback, $U_D = 0$. Where $I_{R1} = I_{R2}$, it is possible in accordance with the voltage divider – consisting of R_1 and R_2 – to calculate the voltage.

$$U_1 = \frac{R_1}{R_1 + R_2} U_A.$$

This means that

$$U_A = \frac{R_1 + R_2}{R_1} U_1 = \left(1 + \frac{R_2}{R_1}\right) U_1.$$

The output voltage U_A is here also directly dependent on the input voltage U_1 and the choice of resistances R_2 and R_1 ; however, here the gain U_A/U_1 has at least the value one; U_A and U_1 are in phase.

Figure 19: Inverting amplifier

U_1 Input voltage,
 U_D Differential voltage,
 U_A Output voltage,
 R_1, R_2 Network resistances,
 I_{R1}, I_{R2} Currents in R_1 and R_2 ,
 I_N Input current.

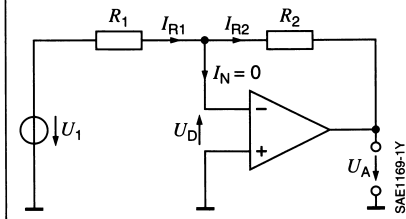
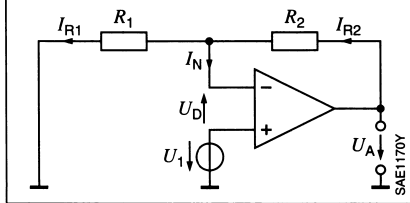


Figure 20: Non-inverting amplifier

U_1 Input voltage,
 U_D Differential voltage,
 U_A Output voltage,
 R_1, R_2 Network resistances,
 I_{R1}, I_{R2} Currents in R_1 and R_2 ,
 I_N Input current.



A special case of the non-inverting amplifier is the isolation amplifier or impedance transformer. If R_1 assumes an infinitely high value (open loop) and R_2 is set equal to zero (short circuit) (Figure 21), then the gain is equal to one (i.e. $U_A = U_1$).

The advantage of this circuit is the property to the effect that the input-voltage source U_1 is not loaded with the internal resistance R_E , as the input current I_P is approximately equal to zero. This results in a negligible voltage drop over R_E and because $U_D = 0$, the input voltage U_1 is available at the output of the operational amplifier as U_A . This is important particularly for conditioning sensor signals, as here the sensor output voltage may not be loaded in many cases, i.e. any current flow from the sensor element can reduce the tappable voltage significantly.

Subtracting amplifier

The subtracting amplifier (Figure 22) can be considered as a common variant of the two previously mentioned circuits. On account of the superposition theorem (superposition principle), the relationship between the output voltage U_A and the input voltages U_1 and U_2 can be derived.

$$U_A = \frac{R_2}{R_1} (U_2 - U_1).$$

Instrument amplifier

Particularly in sensor systems, differential voltages must often be picked off at bridge circuits and amplified without the sensor voltage or the bridge voltage being subjected to an unacceptably high load. This can be implemented by means of a high-resistance voltage tap. An instrument amplifier, which outputs the difference between two potentials U_2 and U_1 as amplified output voltage U_A , can be used for this purpose. The instrument amplifier can be subdivided into two parts: the pre-amplifier and a subtracting amplifier (Figure 22) with further amplification. Figure 23 shows the basic input circuit for the pre-amplification of an instrument amplifier.

According to the rule of negative feedback, the potential difference between the inverting and non-inverting inputs is equal to zero. In each case, the current I can flow through the resistors R and R' , as the input currents I_{N1} and I_{N2} can be ignored. The following applies:

$$I = \frac{U_1 - U_2}{R'} = \frac{U_{A1} - U_{A2}}{2R + R'}, \text{ thus}$$

$$U_{A1} - U_{A2} = (U_1 - U_2) \left(\frac{2R}{R'} + 1 \right).$$

Figure 21: Impedance transformer or isolation amplifier

U_1 Input voltage,
 U_D Differential voltage,
 U_A Output voltage,
 R_E Input resistance,
 I_P Input current.

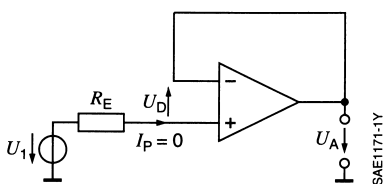
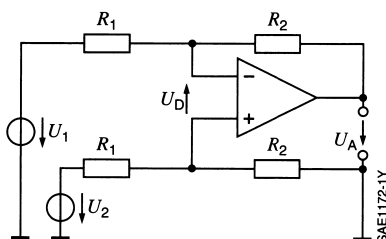


Figure 22: Subtracting amplifier

U_1, U_2 Input voltages,
 U_D Differential voltage,
 U_A Output voltage,
 R_1, R_2 Network resistances.



In this way, the amplified difference between the two voltages U_1 and U_2 is obtained as the potential difference U_D between the two outputs of the two operational amplifiers. In order to output this voltage U_D as an output voltage referred to ground U_A , a subtracting amplifier can be connected in series (Figure 22), where U_{A1} instead of U_1 and U_{A2} instead of U_2 are supplied.

Important characteristic data

For many applications, the operational amplifiers must demonstrate properties which partly contradict each other. There are a large number of operational amplifiers, which are optimized for different fields of application. Basically, the data for specific operating points or operating ranges are specified.

Temperature range

In the field of consumer electronics, a temperature range of between 0 °C and 70 °C is customary. For the extended industrial field, a temperature range of between -20 °C and +70 °C is often specified; this range is required above all for devices which are used outside buildings. For military applications, a temperature range of -55 °C to +125 °C is specified. These demands do not, however, cover all the requirements for use in motor ve-

hicles; for example, even higher temperatures occur in the engine compartment or in brake systems.

Offset voltage

The characteristic quantity for describing the fact that, even in the event of a short circuit between the two inputs (i.e. for $U_D = 0$), the output voltage U_A is not equal to zero is termed the offset voltage U_{OS} . This thus acts like a voltage applied from an external source U_D and is added to this. The offset voltage U_{OS} can, for example, be determined by the fact that at the input that voltage which sets the output voltage U_A equal to zero is determined (Figure 24). The offset voltage U_{OS} stems, among other things, from asymmetries in the internal circuit arrangement of the two inputs and is typically in the range of a few μV to a few mV.

However, as well as the value of the offset voltage U_{OS} , the temperature effect and long-term stability are also highly important. Some operational amplifiers offer the possibility of compensating the offset voltage by means of an external circuit arrangement – provided this is not already effected by internal circuit-engineering measures. It is important in this connection that the input voltage can also drift on account of the temperature effect; thus soldered junctions can represent thermocouples with a voltage of a magnitude of 10 to 100 $\mu V/K$.

Figure 23: Pre-amplification of an instrument amplifier

U_1, U_2 Input voltages,
 I_{N1}, I_{N2} Input currents, I Current
 R, R' Resistances
 U_{A1}, U_{A2} Output voltages, in relation to ground.

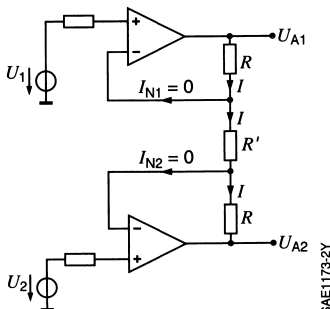
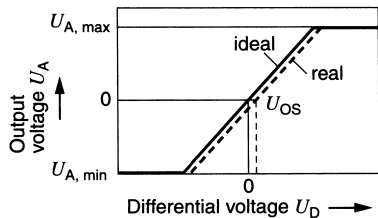


Figure 24: Offset voltage

U_A Output voltage,
 U_D Differential voltage between the inputs,
 U_{OS} Offset voltage,
 $U_{A, MAX}$ Maximum output voltage,
 $U_{A, MIN}$ Minimum output voltage.



Input resistances and currents

On account of the, as a rule, very low input currents I_N and I_P , very high input resistances, which are sometimes in the high megaohm range, are accordingly obtained. A distinction is made here between the common-mode input resistance (resistance between an input in each case and ground) and the differential input resistance between the two inputs.

The inputs of customary operational amplifiers form transistors; either bipolar transistors in which the base in each case is activated or MOS field-effect transistors in which the gate is recharged. The small input currents are explained in this way. When bipolar transistors are used, these are base currents and are in the range of μA . When MOSFET are used, the corresponding gate currents are obtained which are required to recharge the associated gate capacity. The latter are proportional to the operating frequency and are usually in the range of pA.

The input bias current can cause an input-voltage error in high-resistance circuits. This can be compensated for by connecting identical impedances to the two inputs, as then in each case the same voltage drops and the differential voltage U_D remains unaffected. Like the offset voltage, the input current can also drift over the temperature and the time.

Output resistance

The output of the operational amplifier can be described by a series connection of an ideal voltage source and a resistance; the latter is then the output resistance R_A . This resistance limits the output current. In general, operational amplifiers can drive output currents of 20 mA; there are also types with an output current of up to 10 A.

Slew rate

The slew rate (SR) denotes the maximum possible change in the output voltage U_A per period of time, i.e. the maximum value for dU_A/dt . The values for the slew rate for conventional operational amplifiers are in the range of less than 1 V/ μs to more than 1 V/ns.

Noise

Noise can be described by specifying the noise-voltage density or the noise-current density. Usually the noise-voltage density U_R' is given in $\text{nV}/\sqrt{\text{Hz}}$.

The effective value of the noise voltage U_R (this also applies to the noise current) is obtained from the respective key figure multiplied by the root of the bandwidth B being considered:

$$U_R = U_R' \sqrt{B} .$$

For an amplifier circuit, the total effective noise-voltage density is obtained as the root of the sum of the squares of the effective values.

$$U_{R,\text{tot}}' = \sqrt{(U_{R,1})^2 + m (U_{R,m})^2} .$$

m here denotes the number of noise terms.

The noise is determined predominantly at the input of the operational amplifier. If JFET or MOSFET are used, low current but comparatively high voltage noise is obtained. The behavior is reversed in operational amplifiers which are based on bipolar transistors (see [1] and [4]).

Push-pull drivers

Operational amplifiers usually exhibit a maximum output current of approximately 20 mA. If a larger current is needed for example to activate power components or actuators (e.g. DC electric motors), this can be achieved with a driver stage – stages of this kind are also known as push-pull circuits, push-pull stages, or push-pull drivers. It is important in this connection that in some cases a large current I_{load} flows through the load (ohmic resistance R_{load} , Figure 25) in both directions. This is made possible by the use of two power transistors (in Figure 25 T_1 as *npn* and T_2 as *pnp* transistor); MOSFETs can also be used instead of bipolar transistors.

Because of the negative feedback the output voltage U_{OUT} is always equal to U_{IN} (Figure 26) and consequently for an ohmic load the output current I_{load} follows the input voltage U_{IN} .

The two transistors T_1 and T_2 in Figure 25 are activated in such a way that they never conduct simultaneously (push-pull operation) because this could lead to

a short-circuit of the supply voltages $+U_V$ and $-U_V$. Since when the current I_{LOAD} is reversed through the load resistance R_{LOAD} first one transistor and then immediately afterwards the other transistor must conduct, this results in the voltage U_{OPV} in the circuit according to Figure 25 jumping abruptly from for example -0.7 V to $+0.7$ V, as shown in Figure 26. This can give rise to problems in applications with high dynamic response.

Push-pull circuits are frequently also used in driver circuits for IGBTs (insulated-gate bipolar transistors); a positive voltage (frequently $+15$ V) and a negative voltage (frequently -15 V) are made available here for the activation of IGBTs. The driver output stage is as a rule supplied with a metallically separated voltage via an insulated DC/DC converter. In addition, the driver circuit takes on various protective functions such as overcurrent detection or even interlocking logic.

Figure 25: Push-pull stage

Fundamental structure with two bipolar transistors.

- T_1 *nnp* bipolar transistor,
- T_2 *pnp* bipolar transistor,
- OVP_1 Operational amplifier,
- R_{IN} Input resistance,
- R_{LOAD} Load resistance.
- U_{IN} Input voltage,
- U_{OUT} Output voltage,
- I_{LOAD} Load current,
- $+U_V$ Voltage supply, positive potential,
- $-U_V$ Voltage supply, negative potential.

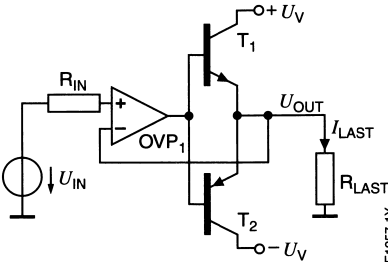
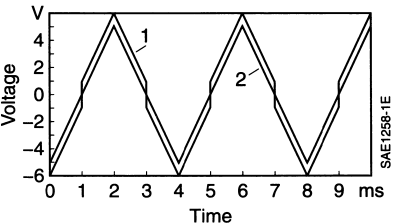


Figure 26: Characteristic curves of a push-pull stage

- 1 Voltage at output of OVP_1 ,
- 2 Output voltage ($U_{OUT} = U_{IN}$).



Monolithic integrated circuits

Monolithic integration

In monolithic integrated circuits (IC), components are assembled on a single piece of monocrystalline silicon (substrate). Semiconductor processes are used for creating layers (e.g. epitaxy), removing layers and changing material properties (e.g. doping). This technology enables complex circuits to be housed in the smallest of spaces.

Planar technology is based on the oxidation of silicon wafers, which is a relatively simple process, and the speed at which the dopants penetrate into silicon, which is exponentially greater than the speed at which they enter the oxide. Doping only occurs at locations where openings are present in the oxide layer. The specific design requirements of an individual integrated circuit determine the precise geometric configuration, which is applied to the wafer in a photolithographic process. All processing procedures (oxidizing, etching, doping, and depositing) progress consecutively from the surface plane (planar).

Planar technology makes it possible to manufacture all circuit components (e.g. resistors, capacitors, diodes, transistors) and the associated conductor strips on a single silicon chip in a unified manufacturing process. Monolithic integrated circuits are built from semiconductor components.

This integration generally comprises a subsystem within the electronic circuit and increasingly comprises the entire system (a system on a chip).

Because of the ever-increasing component density (integration density), the third dimension, i.e. the plane vertical to the surface, is also increasingly being utilized in the design. In this way, particularly in the field of power electronics, advantages such as lower resistances, smaller losses and thus also higher current densities can be achieved.

Integration level

The integration level is a measure of the number of function elements per chip. The following technologies relate to the level of integration (and chip surface):

- SSI (Small Scale Integration) with up to several hundred function elements per chip and a mean chip surface area of 1 mm². But the chip surface area can be very much larger in circuits with high power outputs (e.g. smart power transistors).
- MSI (Medium Scale Integration) with a several hundred to 10,000 function elements per chip and a mean chip surface area of 8 mm².
- LSI (Large Scale Integration) with up to 100,000 function elements per chip and a mean chip surface area of 20 mm².
- VLSI (Very Large Scale Integration) with up to 1 million function elements per chip and a mean chip surface area of 30 mm².
- ULSI (Ultra Large Scale Integration) with over 1 million function elements per chip (flash memory today contains up to 20 billion transistors per chip), a surface area of up to 300 mm² and the smallest structure sizes of less than 30 nm.

Computer-aided simulation and design methods (CAE and CAD) are essential elements in the manufacture of integrated circuits. Entire function modules are used in VLSI and ULSI, otherwise the time expenditure and failure risk would make development impossible. In addition, simulation programs are used to detect any defects made.

References

- [1] U. Tietze, Ch. Schenk, E. Gamm: Halbleiter-Schaltungstechnik. 15th Edition, Verlag Springer Vieweg, 2016.
- [2] A. Führer, K. Heidemann, W. Nerretter: Grundgebiete der Elektrotechnik, Volumes 1–2. 9th Ed., Carl Hanser Verlag, 2011.
- [3] R. Ose: Elektrotechnik für Ingenieure. 5th Edition, Carl Hanser Verlag, 2013.
- [4] R. Müller: Rauschen. 2nd Ed., Springer-Verlag, 2013.

Chemistry

Elements

Periodic table

Structure of the periodic table

The atoms of the chemical elements are made up of positively charged protons, uncharged neutrons, and negatively charged electrons [1]. In the periodic table (Tables 1 and 2) the elements are laid out in order of increasing number of protons – i.e. increasing nuclear charge number – and in order of their atomic mass. The atomic mass is essentially determined by the total number of nuclear constituents, i.e. the sum total of protons and neutrons. The total number of protons and neutrons is also called the mass number. The number of protons corresponds to the atomic number. The neutral element atoms always have as many protons and electrons.

Groups and periods

The elements are divided in the periodic table into different groups (vertical columns) and periods (horizontal rows). The nested structure of the groups is due to the fact that the electrons always occupy the lowest energy levels. The position of these energy levels, which are also called electron orbitals and indicate the probability of electrons being present around the atomic nucleus, can be derived using quantum mechanics.

Quantum numbers

The structure of the periodic table is based on four quantum numbers (principal, secondary, magnetic, and spin quantum numbers), with which the four electron orbitals with the designations s, p, d and f can be calculated. When the electrons are arranged into these orbitals in order of increasing energy, this creates groups of elements which demonstrate a similar manner of reaction across all the periods. This manner of reaction is only minimally influenced by the electrons on the inner orbitals. The crucial factors are the energy and number of outer electrons. The outer electrons are frequently referred to as "valence electrons".

Main groups

The elements in groups Ia, IIa and IIIa...VIIa are called main-group elements. The main-group elements include hydrogen and the alkali metals (Ia), the alkaline-earth metals (IIa), the boron (IIIa), carbon (IVa) and nitrogen groups (Va), the chalcogens (VIa), the halogens (VIIa), and the noble gases (VIIIa). The electrons of the main-group elements of the 1st period, hydrogen and helium, are found exclusively in s-orbitals. In the case of the other main-group elements in the 2nd through 7th periods, the electrons also occupy p-orbitals starting from the main group IIIa.

Secondary groups

The elements of the secondary groups Ib, IIb and IIIb...VIIIb with electrons in d-orbitals all have a metallic character. The copper group is assigned to the secondary group Ib because its electron configuration shows similarities to the main group Ia. The elements in both groups tend to form salts from monovalent ions. The same applies to the groups IIa and IIb. Divalent metal compounds are formed in the group of alkaline-earth metals and the zinc group IIb. The designations for the secondary groups IIIb...VIIIb likewise have their roots in the electron configuration and provide an indication of the maximum valence of these metal ions. The elements iron, cobalt and nickel as well as the elements underneath called higher homologs are brought together as secondary group VIIIb due to their marked chemical similarity.

Table 1: Periodic table of elements

Ia	IIa																IIIa			IVa			Va			Via			VIIa			VIIa
	1	IIa															IIIa			IVa			Va			Via			VIIa			2
	H																B	C	N	O	F							He				
1.008																10.811	12.011	14.007	15.999	18.998										4.003		
3	4																															
Li	Be																											10				
6.941	9.012																											Ne				
11	12																											18				
Na	Mg																											Ar				
22.990	24.305															26.982			30.974			32.066			35.453			39.948				
19	20															31			32			33			34			36				
K	Ca															Ga			Ge			As			Se			Kr				
39.098	40.078															69.723			72.61			74.922			78.96			79.904			83.80	
37	38															47			48			51			52			54				
Rb	Sr															Ag			Cd			Sb			I			Xe				
85.468	87.62															107.868			112.411			121.760			127.60			126.904			131.29	
55	56															81			80			83			84			86				
Cs	Ba															Tl			Pb			Bi			Po			Rn				
132.905	137.327															200.59			204.383			206.980			(209)			(210)			(222)	
87	88															112			111			115			116			118				
Fr	Ra															Cn			Nh			Mc			Lv			Ts				
(223)	(226)															(285)			(284)			(291)			(293)			(292)	(294)			

All elements are arranged sequentially according to atomic number (proton number). The horizontal rows are called periods, and the vertical columns are called groups. The relative atomic masses are indicated below the element symbols. The values given in parentheses are the mass numbers (nucleon numbers) of the stablest isotopes of radioactive elements.

Table 2: Designations of chemical elements

Element	Sym- bol	Atomic number	Element	Sym- bol	Atomic number	Element	Sym- bol	Atomic number
Actinium	Ac	89	Indium	In	49	Radium	Ra	88
Aluminum	Al	13	Iodine	I	53	Radon	Rn	86
Americium ¹	Am	95	Iridium	Ir	77	Rhenium	Re	75
Antimony	Sb	51	Iron	Fe	26	Rhodium	Rh	45
Argon	Ar	18	Krypton	Kr	36	Roentgenium ¹	Rg	111
Arsenic	As	33	Lanthanum	La	57	Rubidium	Rb	37
Astatine	At	85	Lawrencium ¹	Lr	103	Ruthenium	Ru	44
Barium	Ba	56	Lead	Pb	82	Rutherfordium ¹	Rf	104
Berkelium ¹	Bk	97	Lithium	Li	3	Samarium	Sm	62
Beryllium	Be	4	Livermorium ¹	Lv	116	Scandium	Sc	21
Bismuth	Bi	83	Lutetium	Lu	71	Seaborgium ¹	Sg	106
Bohrium ¹	Bh	107	Magnesium	Mg	12	Selenium	Se	34
Boron	B	5	Manganese	Mn	25	Silicon	Si	14
Bromine	Br	35	Meitnerium	Mt	109	Silver	Ag	47
Cadmium	Cd	48	Mendelevium ¹	Md	101	Sodium	Na	11
Carbon	C	6	Mercury	Hg	80	Strontium	Sr	38
Cesium	Cs	55	Molybdenum	Mo	42	Sulfur	S	16
Calcium	Ca	20	Moscovium ¹	Mc	115	Tantalum	Ta	73
Californium ¹	Cf	98	Neodymium	Nd	60	Technetium	Tc	43
Cerium	Ce	58	Neon	Ne	10	Tellurium	Te	52
Chlorine	Cl	17	Neptunium ¹	Np	93	Tennessine ¹	Ts	117
Chromium	Cr	24	Nickel	Ni	28	Terbium	Tb	65
Copernicium ¹	Cn	112	Nihonium ¹	Nh	113	Thallium	Tl	81
Copper	Cu	29	Niobium	Nb	41	Thorium	Th	90
Cobalt	Co	27	Nitrogen	N	7	Thulium	Tm	69
Curium ¹	Cm	96	Nobelium ¹	No	102	Tin	Sn	50
Darmstadtium ¹	Ds	110	Oganesson ¹	Og	118	Titanium	Ti	22
Dubnium ¹	Db	105	Osmium	Os	76	Tungsten	W	74
Dysprosium	Dy	66	Oxygen	O	8	Uranium	U	92
Einsteinium ¹	Es	99	Palladium	Pd	46	Vanadium	V	23
Erbium	Er	68	Phosphorus	P	15	Xenon	Xe	54
Europium	Eu	63	Platinum	Pt	78	Ytterbium	Yb	70
Fermium ¹	Fm	100	Plutonium ¹	Pu	94	Yttrium	Y	39
Fluorine	F	9	Polonium	Po	84	Zinc	Zn	30
Flerovium ¹	Fl	114	Potassium	K	19	Zirconium	Zr	40
Francium	Fr	87	Praseodymium	Pr	59			
Gadolinium	Gd	64	Promethium	Pm	61			
Gallium	Ga	31	Protactinium	Pa	91			
Germanium	Ge	32						
Gold	Au	79						
Hafnium	Hf	72						
Hassium ¹	Hs	108						
Helium	He	2						
Holmium	Ho	67						
Hydrogen	H	1						

¹ Artificially produced; does not occur naturally.

Periods with f-orbitals

In the 6th and 7th periods a further series of energy levels, the f-orbitals, are available in each case after the secondary group IIIb lanthanum and actinium. These are taken up by the electrons of the elements of the lanthanides (6th period) and actinides (7th period). The lanthanides are also known by the term "rare earths". All actinides are radioactive.

The arrangement of the elements in the periodic table according to the energy position of their electron orbitals becomes clear in the energy-level diagram (Figure 1). It can be seen that for higher orbital energies after the occupation of the s-orbital the p-, d- and f-orbitals subordinate to this principal quantum number are no longer automatically filled with electrons. For example, it is now understandable why the secondary-group elements scandium through zinc with the orbital energies 3d after the element calcium are placed into the 4th period of the periodic table. For higher nuclear charge numbers, however, the differences between the orbital energies are so small that the energy sequence of the orbitals shown in the energy-level diagram no longer applies to each individual element in the periodic table. In these cases the exact energy position is influenced by the partial, semi-

or full occupation of the orbitals by electrons. This can be seen in the case of the lanthanides with the 4f-orbitals, which are only placed after the element lanthanum into the 6th period and not – as would be expected according to the energy-level diagram – already after the element barium.

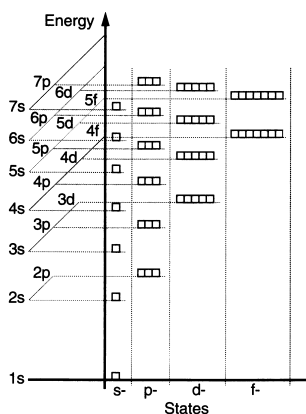
Isotopes

Isotopes are atoms of the same element with the same number of protons but with a different mass number, i.e. with a different number of neutrons. To give isotopes a unique designation, the mass number is given at top left next to the atom symbol and the number of protons is given at bottom left so that the number of neutrons can be calculated straight away by subtraction. Most naturally occurring elements are isotopic mixtures. Carbon, for example, is made up 98.89 % of $^{12}_6\text{C}$ and 1.11 % of $^{13}_6\text{C}$. The proportion of $^{14}_6\text{C}$ is at $10^{-10}\%$ very low. $^{14}_6\text{C}$ in comparison with $^{12}_6\text{C}$ and $^{13}_6\text{C}$ not a stable isotope. The ratio of $^{14}_6\text{C}$ to $^{12}_6\text{C}$ is used in the radiocarbon method to determine the age of organic matter.

Nuclides

The term nuclide is broadly defined. All atoms that differ in the constitution of their nucleus are called nuclides. The number of nuclides therefore corresponds to the number atom types. The 8th Edition of the Karlsruhe Nuclide Chart in 2012 lists 3,847 experimentally verified nuclides and isomers. Isomers are nuclides which have the same mass number and thus the same number of protons and neutrons. Isomeric nuclides, however, show differences in the internal state of the atomic nuclei. Atomic nuclei can assume excited states as well as the normal state. Only around 270 nuclides are stable. The majority of nuclides are radioactive and are therefore called radionuclides.

Figure 1: Energy-level diagram of the electron orbitals



UAM0239-2E

Radioactive decay *α -decay*

In the case of natural radioactivity [2] α -decay is observed when the atomic nucleus emits a double positively charged helium nucleus He^{2+} consisting of two protons and two neutrons. The mass number of the atomic nucleus is reduced

by 4 and the nuclear charge number by 2. A different element is created (example, see Figure 2a).

 β -decay

An element transmutation also occurs in the case of radioactive β -decay. β^- decay is observed when a neutron from the atomic nucleus emits an electron (example, see Figure 2b). This increases the number of protons by 1 and the element is created. When instead of an electron a positron, i.e. a positively charged particle, is emitted, β^+ -decay takes place. A proton is transformed into a neutron in the case of β^+ -decay. Where the mass number is the same, the nuclear charge number is therefore decreased by 1 such that the preceding element in the periodic table is created (example, see Figure 2c). β -decay also witnesses the release of antineutrinos (in the case of β^- -decay) or neutrinos (in the case of β^+ -decay), which cannot be discussed further at this stage.

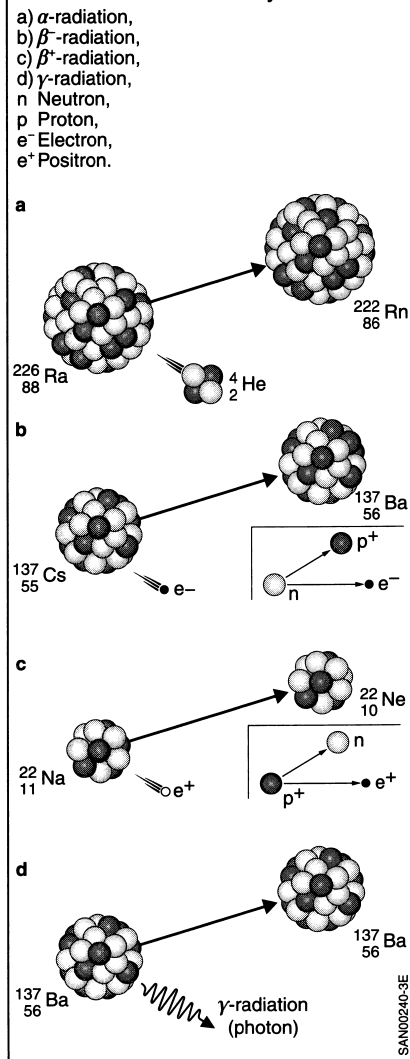
 γ -decay

In the case of γ -decay the atomic nucleus emits internal energy from excited nucleus states which for the most part arise during α - or β -decay. The γ -radiation emitted during γ -decay does not change the mass number; the element remains intact (example, see Figure 2d).

Artificial element transmutation

Elements can also be artificially transmuted into each other by bombardment with high-energy particles. Neutrons, singly or doubly positively charged hydrogen (H^+) and helium nuclei (He^{2+}), but also γ -radiation are suitable for this purpose. Depending on the type and acceleration of the particles, this results in the particle being absorbed in the atomic nucleus – sometimes also involving the emission of a proton or neutron – or in nuclear fission.

Figure 2: Types of radiation for natural radioactive decay

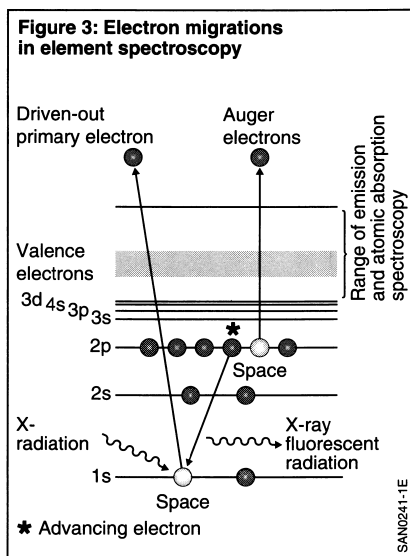


Half-life

Radioactive decay is a monomolecular reaction (first-order reaction) in which the decaying share per unit of time is directly proportional to the available quantity. Because the available quantity constantly diminishes due to the radioactive decay continuously taking place, the reaction rate decreases more and more. The half-life of the radionuclide is specified so that the rate of radioactive decay can be described independently of the quantity. The half-life is the period of time after which half the atomic nuclei have decayed. The shorter the half-life of the radioactive element, the higher its specific activity, i.e. the more nuclear decays occur per unit of mass.

Element spectroscopy

All chemical elements differ in their atomic structure and can therefore be identified by means of their electron spectra and quantified into mixtures [3]. To release electrons from their energy levels, it is necessary to supply energy to the order of their binding energy from an outside source.

**X-ray fluorescence analysis**

The inner electrons more tightly bound by the attraction of the positive atomic nucleus can only be released with X-radiation or with electrons of sufficient energy. X-ray fluorescence analysis (XRF) is based on the fact that the spaces created on the more favorable inner energy levels are filled by higher-energy electrons. These advancing electrons release the excess energy in the form of fluorescent radiation. The energy of the fluorescent radiation corresponds exactly to the difference between the higher and the lower energy levels and is thus element-specific (Figure 3).

Auger electron spectroscopy

The energy released by advancing electrons can be output as fluorescent radiation, but can also be transmitted to other electrons. An electron can leave the atom with the transmitted energy. These electrons from the so-called Auger process likewise contain highly specific element information, since the following contributions determine the energy of the Auger electron: The orbital energy of the originally driven-out electron, the energy that is released by the advancing electron, and the energy of the electron emitted in the Auger process itself (Figure 3). The energy of the emitted electrons is analyzed in Auger electron spectroscopy.

Emission spectrometry

The less tightly bound outer valence electrons can already be thermally excited and thereby raised to higher, unoccupied energy levels. Because the energy from the flame of a gas torch is not sufficient to excite the valence electrons of all the elements, a plasma is often used for excitation. In the case of inductively coupled plasma optical emission spectrometry (ICP-OES), the light released during the fallback of the valence electrons that is in the visible and ultraviolet range is registered. As with the inner electrons, the energy difference between the outer occupied and unoccupied energy levels is also element-specific.

Atomic absorption spectroscopy

Instead of emitted radiation, light absorption can also be measured through the excitation of the valence electrons. To this end, in atomic absorption spectroscopy (AAS) the emitted radiation of the element to be analyzed is irradiated by a hollow-cathode lamp. When the element in the sample matches the element of the hollow cathode, the excitation of the electrons in the sample causes a weakening in the intensity of the excitation light. Because the verification of different elements accordingly requires an extensive range of materials for hollow-cathode lamps and can only be measured sequentially, atomic absorption spectroscopy has not established itself in multi-element analyses.

X-ray photoelectron spectroscopy

In X-ray photoelectron spectroscopy (XPS) electrons are driven out from all the energy levels in the atom by high-energy X-rays. The element-specific binding energy can be inferred from the kinetic energy of these electrons. Only the valence electrons are released by the ultraviolet light used instead of X-rays in ultraviolet photoelectron spectroscopy (UPS). The kinetic energy of valence electrons can be determined very precisely such that even conclusions can be made as to the type of chemical bonds by way of differences in the orbital energies.

X-ray diffraction

X-ray diffraction on solids is, as the name suggests, not a spectroscopic procedure. The X-radiation diffracted on the crystal lattice is registered with angular resolution, i.e. according to its diffraction angle [4]. The angle position, intensity and width of the diffraction peaks enable the lattice structure to be analyzed and thus crystals to be identified (crystal-structure analysis). Crystallite size, preferred orientations (textures) and lattice distortions can also be determined in mixtures. Internal mechanical stresses of components can likewise be determined.

Chemical bonds

The type of bond which tends to be adopted by the individual elements is determined by the number and arrangement of electrons around the atomic nucleus (electron configuration) [5]. The nature of these bonds can differ greatly. A distinction is made between ionic, covalent and metallic bonds. Single atoms without a chemical bond occur in nature in noble gases and in some elements in the vapor phase.

Additional, but far weaker interactions exist between molecules, i.e. particles which are composed of two or more atoms. The weak attractive forces, which have different physical causes, mean that not only ions but also molecules adopt a short-range order, which is only eliminated in the gaseous phase.

Ionic bonds

Ionic bonds tend to form in compounds between metallic and non-metallic elements. This results in the transfer of electrons. Metal atoms (electron donors) lose electrons and become positively charged cations. By gaining electrons, the non-metal atoms (electron acceptors) become negatively charged anions. The terms "cation" for a positively and "anion" for a negatively charged ion are so called because these ions move in an aqueous solution in the electric field to the oppositely charged electrodes, i.e. the cation to the negatively charged cathode and the anion to the positively charged anode.

The number of lost and gained electrons is determined by the electron configuration of the element. Ionic compounds with configurations in which the p-, d- and f-orbitals are empty, half-full and full are especially preferred.

The ionic compounds created are called salts. In a solid the cation and anion are arranged in an ionic lattice whose structure is determined by the ratio of ionic radii. When salts are melted or dissolved in water, the forces between the ions in the lattice must be overcome. Therefore heat of fusion (enthalpy of fusion) must be supplied. The energy required for dissolving (heat of solution) is determined by two opposing components. Firstly, energy to dissolve the crystal lattice must be applied (heat of dissociation); secondly, energy is released by coordination of the solvent molecules to the dissolved-out ions (heat of solvation). If the solvent is water, the heat is called heat of hydration. If the energy of dissociation is greater than the heat of solvation, the solution cools during the dissolving process. When anhydrous metal salts are dissolved in which initially water is also introduced into the ionic lattice, the heat released by hydration is often greater than the energy consumption to overcome the lattice forces, and the solution heats up.

Covalent bonds

Bonds of electron pairs which are created between two neutral atoms from the unpaired electrons in the outer orbitals (valence electrons) are called atomic bonds. The term "covalent" illustrates that two identical or similar atoms can, by forming a common electron pair, achieve an energetically more favorable electron configuration without involving an electron transfer and with it a change in the valency by the formation of ions. The configuration of the noble gas from the period in which the element is found is usually adopted as an energetically more favorable state.

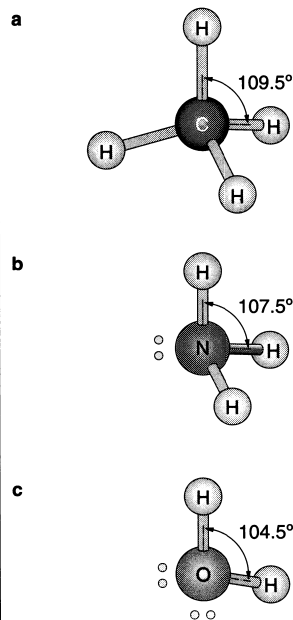
In a methane molecule (CH_4) all the atoms have a noble-gas configuration. Carbon has in the p-orbitals two valence electrons and four free spaces for further electrons which are made available by the four hydrogen atoms. Carbon thus achieves the noble-gas configuration of neon. Because the two atoms share the common bonding-electron pair, the hydrogen atoms, which only have in each case one valence electron in the s-orbital, also attain the noble-gas configuration of helium through the common electron pairs with the carbon. In the case of the main-group elements of the 2nd period, the filling of empty orbitals with valence electrons of other atoms until the noble-gas configuration is attained is called the octet rule, because then a total of eight electrons, two in s- and six in p-orbitals, are present.

An atom can enter into several covalent bonds with a neighboring atom, such as for example carbon double bonds in ethylene ($\text{H}_2\text{C}=\text{CH}_2$) or triple bonds in acetylene ($\text{H}-\text{C}\equiv\text{C}-\text{H}$).

In the octet rule existing free electron pairs which have not bonding function are also taken into account. The three-dimensional structure of molecules is determined by the total number of atoms, by the type and number of bonding-electron pairs, and by the available free electron pairs. For example, methane (CH_4), ammonia (NH_3) and water (H_2O) have a tetrahedral basic structure. However, only the methane molecule represents an ideal tetrahedron with a bond angle of 109.5° (Figure 4). Free electron pairs require a greater amount of space than bonding-electron pairs. The available free electron pair in ammonia reduces the bond angle to 107.5° . The water molecule has two available free electron pairs, further reducing the bond angle to 104.5° .

Figure 4: Influence of free electron pairs on the structure of molecules

- a) Methane (CH_4),
b) Ammonia (NH_3),
c) Water (H_2O).



In the simplified model presented here covalent atomic bonds are put down to the formation of pairs of valence electrons from atomic orbitals. However, solely examining atomic orbitals only touches the surface of the complex conditions in a molecule. All the atomic nuclei and electrons in a molecule influence each other, and the bonding conditions are therefore better described by molecular orbitals. A quantum-mechanical examination of molecular orbitals is required to describe double and triple bonds and delocalized bonding systems in aromatic compounds; it is perfectly sufficient in this case to confine oneself to the valence electrons.

Metallic bonds

In metals the atoms are arranged in three-dimensional lattices where the valence electrons of the atoms move freely and are therefore often referred to as electron gas. Metals are good conductors of electricity and heat, which can be put down to the free valence electrons. The lattice vibrations of the atoms also control to the thermal conductivity; here the atoms vibrate freely in their positions in the metal lattice and can easily transmit heat in this way. Here too, a quantum-mechanical examination of molecular orbitals is required to better understand the greater complexities of the electric conductivity of metals. It is possible to derive from this that the electrons reside in energy bands which are formed from molecular orbitals with the smallest energy differences. Between the energy bands are zones which cannot be occupied by electrons.

Interactions between molecules

Van der Waals forces

Partial charges which are created by three-dimensional fluctuations of the positive and negative charge concentrations in the molecule induce electric dipoles (Figure 5a). Interactions that are created between the molecules polarized in this way are called van der Waals forces. The bigger the molecule, the more it can be polarized and the stronger the intermolecular forces.

Dipole-dipole interactions

In molecules which are composed of different atoms the charge concentration is continuously displaced (Figure 5b). Depending on the size and the charge of the atomic nuclei and inner electrons involved, the valence-electron pair is subjected to different degrees of attraction. The property of atoms to attract the bonding-electron pair is known as electronegativity. The more strongly attracting, more electronegative atom receives a partially negative charge. Correspondingly, the more weakly attracting atom receives a partially positive charge. The permanent dipoles brought about by a difference in the electronegativity exhibit intermolecular forces that are much stronger than van der Waals forces. These phenomena are called dipole-dipole interactions.

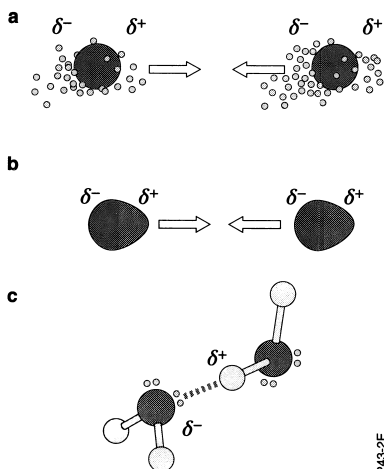
The dipole-dipole interactions are also influenced by the three-dimensional structure of the molecule. Although there are differences in the electronegativity of carbon and oxygen, carbon dioxide (CO_2) for example is not a dipole on account of its linear structure. Water (H_2O) with the same atomic ratio 2:1, however, is a dipole, because the central oxygen atom has two free electron pairs, which result in an angled structure.

Hydrogen bonds

The angled structure of the water molecule facilitates additional dipole-dipole interactions (Figure 5c). The partially positively charged hydrogen atoms can interact with the free electron pairs of neighboring molecules. In liquid water the additional dipole-dipole interactions provide a particular short-range structure, which is responsible for water having the highest density at 4°C . This enables fish to survive in winter because water with the highest density collects at the bottom of frozen lakes.

Figure 5: Types of interactions between molecules

- a) Van der Waals forces,
 b) Dipole-dipole interaction,
 c) Hydrogen bond.
 δ^+ Positive partial charge,
 δ^- Negative partial charge.
 Arrows indicate the force action.



SAND243-2E

Molecular spectroscopy

Molecules, like atoms, can absorb energy and assume defined energy states. Valence electrons can be excited with ultraviolet or visible light (UV/VIS spectroscopy) and vibrations and rotations with infrared light (IR spectroscopy) respectively [6]. These procedures are used to explain structures because energy absorption allows conclusions to be drawn about molecular structure. In Raman spectroscopy the inelastic scattering of monochromatic light of molecules is analyzed. The frequency-shifted portions for example resulting from excitation of vibrations provide structural information.

Mass spectroscopy on the other hand is not based on the excitation of energy states. In this procedure molecules are ionized and the ions obtained are separated according to their mass-to-charge ratio and identified.

Substances

Substance term in chemistry

The chemical properties of a body are determined by its material and not by its size and shape. The material of a body is therefore also called substance. Even where the body has a very fine distribution the chemical properties remain the same. However, a large surface can increase the reactivity, as can be seen in nanoparticles.

Homogeneous and heterogeneous substances

Substances with a uniform structure are termed single-phase or homogeneous. A substance that is made up of two or more parts which cannot be mixed with each other is termed heterogeneous. An example of a homogeneous solid substance is elementary sulfur. Typical of a heterogeneous solid mixture is granite, which is made up of quartz, feldspar and mica.

Dispersion

Heterogeneous mixtures are always dispersions. Dispersions consist of at least two different substances which under the prevailing conditions do not or barely dissolve in each other and do not chemically react with each other. Depending on the phases, a distinction is made between suspensions (liquid and solid), emulsions (liquid and liquid) and aerosols (gas and solid or gas and liquid).

Suspension

If, for example, clay is added to pure water – a homogeneous liquid – this creates a heterogeneous mixture of a liquid and a solid which is termed a suspension.

Emulsion

Heterogeneous mixtures of two liquids, e.g. water and oil, are called emulsions.

Aerosol

A heterogeneous mixture of solid or liquid floating particles and a gas is called an aerosol. An example of an aerosol of gases and solids is exhaust gas with soot particles. Exhaust gas with white smoke, which is produced when water and sulfu-

ric acid condense during the starting operation in the still cold exhaust gas, is on the other hand an aerosol of gases and liquids.

Colloid

The term colloid is used for particles or droplets ranging in size between 1 nm and 1 μm , irrespective of whether the heterogeneous mixtures are suspensions, emulsions or aerosols.

States of aggregation

The three classic states of aggregation (phases) are solid, liquid and gaseous – depending on whether the particles are in a fixed position in a solid body, can move in a liquid while maintaining a short-range order, or are far apart from each other in a gas. Plasma is a non-classic state of aggregation and consists of free electrons and ionized atoms.

The states of aggregation of substances are pressure- and temperature-dependent and are described in the state or phase diagrams. State diagrams, in which pressure is plotted against temperature, explain the conditions under which a solid is present, if necessary in different forms, the so-called modifications. Modifications refer to the phenomenon when a substance in the solid state occurs in different structural forms. Also following from the state diagrams are the ranges in which a liquid or a gaseous phase is present.

State diagram for carbon

The state diagram for carbon shows the solid modifications of graphite and diamond (Figure 6), where in ranges of these phases the other modification exists in parallel as a metastable form. Metastable means that the transformation of a modification is inhibited despite its higher energy content for small changes of state. Graphite is the stable modification at room temperature and atmospheric pressure; diamonds exist however as a metastable structure variant because the transformation into the more stable graphite form is strongly curbed by a high activation energy.

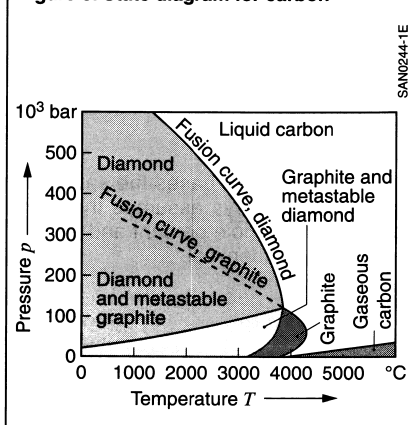
State diagram for water

The three areas in the state diagram for water show the existence ranges of ice, liquid water and vapor (Figure 7). Within these areas only the respective state of aggregation prevails. The curves between the areas describe the equilibrium between the liquid and gaseous phases (vaporization curve), the solid and liquid phases (fusion curve), and the solid and gaseous phases (sublimation curve). Each point on the curves corresponds to a state of equilibrium between the adjacent phases. At the point where the three curves intersect, the triple point TP (0.01°C ; 6.1 mbar), all three phases of water at equilibrium with one another. If the temperature or pressure is changed, only one state of aggregation still exists. If temperature and pressure are simultaneously changed in such a way that the new state corresponds to a point on one of the equilibrium curves, two phases exist next to each other.

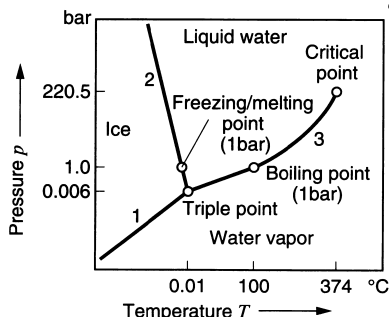
An increase in temperature always results in a higher proportion of gaseous water, either through greater sublimation or through a higher vapor pressure. Conversely, a higher ambient pressure results in more condensation or resublimation of water vapor. The vapor-pressure curve demonstrates that the boiling point of water is dependent on external pressure. At low air pressure or in a vacuum water boils at temperatures below 100°C .

What is conspicuous is the anomaly that can be observed at the phase transition from solid to liquid. An increase in both temperature and pressure causes ice to liquify. The transformation of ice into liquid water is furthered by pressure, because water assumes a smaller volume than ice.

A special thermodynamic state is the critical point CP (374°C ; 220.5 bar), at which the densities of the gaseous and liquid phases become indistinguishable. At the end of the vaporization curve the two states of aggregation – fluid and gaseous – pass into a new state, the supercritical phase. Water in the supercritical state is a liquid which has a lower density than liquid water below the critical point.

Figure 6: State diagram for carbon**Figure 7: State diagram for water**

- 1 Sublimation curve,
- 2 Fusion curve,
- 3 Vaporization curve.



Substance concentrations

To be able to describe the quantitative proportion of the substances involved in a chemical reaction, it is necessary to have quantitative details which refer to the number of particles involved in the conversion.

Owing to the different structure of atomic nuclei all element atoms and naturally also all the resulting molecular compounds differ in mass. It would not be very practicable to reckon with the minute masses of individual atoms or molecules directly. Because it is also helpful for chemical reactions always to consider the same number of particles, the term amount of substance was introduced with the unit "mole".

Amount of substance

An amount of substance n of 1 mole always contains the same number of particles and was therefore previously also often referred to as mole number. Irrespective of the chemical element, molecule or type of substance, 1 mole contains approximately $6 \cdot 10^{23}$ particles. This particle number can be experimentally determined and is also known as "Avogadro's number".

Molar mass

1 mole of a substance therefore always consists of around $6 \cdot 10^{23}$ particles. For element atoms the mass of these particles (molar mass M) therefore corresponds to their relative atomic mass, e.g. 1 mole of carbon has a mass of 12.011 g.

For chemical compounds the molar mass is calculated from the molar masses of the contained elements. Carbon dioxide (CO_2) has a molar mass of 44.01 g/mole (12.011 g/mole for carbon and $2 \cdot 15.9994$ g/mole for oxygen).

The molar mass of a substance – the mass of 1 mole of particles – therefore corresponds to a macroscopically manipulable amount of substance. The molar mass M is the quotient of the mass m and the amount of substance n of a substance (unit g/mol):

$$M = \frac{m}{n}.$$

In chemistry the term weight is also used instead of mass. Mass denotes the matter contained, weight the force acting as a result of the gravitational field on this matter. Since the force of gravity on Earth is approximately always the same, often no distinction is made between mass and weight. That is why the term "molar weight" is also widely used.

Molar volume

Since 1 mole of a substance always consists of the same number of particles, the volume (molar volume V_M) that these particles assume is always the same – provided the particles do not influence each other. This limiting case is only given for an ideal gas: 1 mole of an ideal gas assumes under normal pressure ($p = 1013.25$ mbar) at $T = 273.15$ K (0°C) a volume of 22.414 l and at $T = 298.15$ K (25°C) a volume of 24.789 l. With regard to the standard conditions, weight- and volume-specific concentration details about the molar volume can be approximately converted into each other for all gases.

Mole percent

Some concentration details are given in mole percent. A distinction is made here between the amount-of-substance- and volume-related definitions. The amount-of-substance-related concentration mole percent x with the unit % (n/n) is obtained by multiplying the mole fraction x_i – i.e. the amount of substance n_i of the constituent i referred to the sum of the amounts of substance of all the constituents of the substance mixture – by 100 %:

$$x = \frac{n_i}{\sum_{j=1}^n n_j} \cdot 100.$$

In the case of ideal gases the same particle number always assumes the same volume so that mole percent and percent by volume are identical.

Reactions of substances

Chemical thermodynamics

In chemical reactions starting substances are converted into reaction products with different properties. Chemical thermodynamics describes the substance conversion and the accompanying change in the internal energy ΔU , i.e. in the end whether and under what conditions reactions can take place [7]. The reaction path is determined by the amount of absorbed or output energy.

Heat of reaction

Most reactions are determined in open vessels, i.e. at a constant air pressure. At a constant pressure the change in the internal energy in a chemical reaction is made up of the two components heat of reaction Q_P and work W :

$$\Delta U = Q_P + W.$$

Mechanical work is done for example in an exothermic reaction in which a gas is produced which expands against atmospheric pressure or presses against a membrane or a moving punch.

Reactions can also be effected under a constant volume. Because then the energy content can only be changed by the heat of reaction Q_V and not by work W , the change in the internal energy ΔU corresponds to the heat of reaction Q_V . The following applies:

$$\Delta U = Q_V.$$

The heat of reaction at constant volume Q_V is therefore always greater than Q_P at constant pressure.

Enthalpy of reaction

The proportion of the heat of reaction Q_P or Q_V can also be described as the difference in heat contents of reaction and starting products and is called the enthalpy of reaction ΔH_R . Where heat is absorbed, the reaction is endothermic ($\Delta H_R > 0$). Chemical reactions in which heat of reaction is released ($\Delta H_R < 0$) are exothermic.

The relationship

$$\Delta U^0 = \Delta H_R^0 + W^0$$

is known as the first law of thermodynamics. As the change in the internal energy ΔU , the enthalpy ΔH_R and the work done W are dependent on the pressure and temperature, the standard conditions at 25 °C and 1,013 mbar are referenced and this is indicated at the symbols with a superscript zero (e.g. for ΔU^0).

Enthalpy of activation

Even exothermic reactions often have an initial energy demand despite the released enthalpy of reaction ΔH_R . This enthalpy of activation ΔH_A must be expended in order for example to break bonds and establish an activated complex of reacting agents before more energy is released by the creation of new bonds during the formation of the reaction product than was required to break the bonds in the starting products.

Thermodynamically and kinetically controlled reaction progression

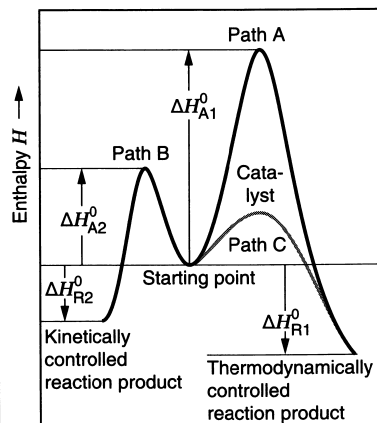
The thermodynamically stable, i.e. energetically more favorable reaction product is reproduced in a reaction when the enthalpy of activation required for substance conversion is applied, e.g. through the supply of heat from the outside (Figure 8, path A). However, if the enthalpy of activation ΔH_{A1}^0 which is required to create the thermodynamically more stable product is not available, another chemical reaction may nevertheless under certain circumstances occur, namely the creation of the product kinetically controlled by the reaction rate. To create it an enthalpy of activation ΔH_{A2}^0 is likewise required (path B), but this is lower in terms of absolute amount. This product created by kinetic control of the reaction then has a higher internal energy. By choosing the reaction conditions, i.e. for example via the quantity of heat made available to the reaction, it is possible to influence whether the thermodynamically or kinetically controlled product tends to be created.

Catalyst decrease the enthalpy of activation of a reaction and thereby increase the reaction rate; but they cannot shift the thermodynamic equilibrium (path C).

Entropy

The enthalpy of reaction ΔH_R alone is not the only factor that determines the direction of a reaction. The distribution of the energy content also plays a role. The evaporation of a liquid takes place despite the required heat of evaporation $\Delta H_V > 0$ because the constraints for the individual molecules decrease as a result. The kinetic energy distributed to the individual molecules passes over from a more ordered, less probable distribution to a state of lesser order but greater probability. The change in the state of order is described by way of the characteristic entropy S [7]. Entropy increases in chemical reactions when more molecules are created than abreact, when the temperature rises, or when particles change to a less ordered state of aggregation (e.g. from solid to liquid or from liquid to gaseous). The increase in entropy ΔS is greater, the higher the heat input and the lower the temperature T at which the heat is transferred.

Figure 8: Enthalpy changes in thermodynamically and kinetically controlled reactions



SAN0246-1E

Reaction kinetics

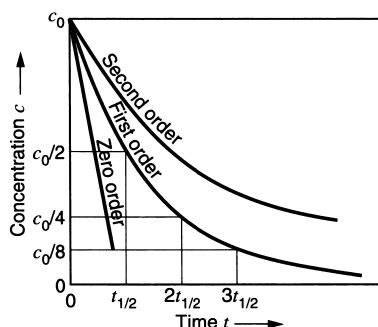
Whereas thermodynamics describes the conversion of substance and energy, reaction kinetics is concerned with the velocity or rate of reactions [7]. The reaction rate of a substances involved in the reaction is specified as a change in concentration per time interval (Figure 9). The order of a reaction is derived from the number of starting substances whose concentration changes during the reaction.

Zero-order reaction

A zero-order reaction exists for example when the decay of a gas on a platinum surface is heterogeneously catalyzed. The concentration of the gas adsorbed on the catalyst does not change during the reaction such that the reaction rate always remains the same, regardless of the reaction time.

Figure 9: Decrease in the concentration of a starting substance in reactions of different orders (assuming equal starting rates)

$t_{1/2}$ Half-life.



SAN0247-2E

First-order reaction

In a first-order reaction the reaction rate depends only on the concentration of the or a starting product. Many decay processes such as radioactive decay for example follow a first-order reaction which is characterized by the fact that the time in which the concentration of the starting product decreases by half (half-life) is always the same. The half-life is therefore not dependent on the starting concentration.

Second-order reaction

A reaction between two molecules is called a second-order reaction when the reaction rate is determined by the concentration of the two starting substances. If one of the reactands is present in a very large surplus, as is the case for example in hydrolysis reactions in which water is simultaneously the solvent and the reacting agent, the reaction rate equates to that of a first-order reaction. This is called a pseudo-first-order reaction.

With equal starting concentration and starting rate a second-order reaction always runs more slowly than a first-order reaction. The rate of the two reaction types is dependent on the concentration of the starting substances. In a second-order reaction, however, it is necessary to factor in that for a substance transformation – unlike in a first-order reaction – a collision of molecules of the two reacting agents is required and also not every collision actually results in a transformation.

Third-order reaction

The probability of a trimolecular reaction, i.e. the simultaneous collision of three particles, is statistically low. For this reason, third-order reactions occur only very rarely.

Chemical equilibrium

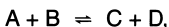
Most chemical reactions are reversible, i.e. an equilibrium is obtained between the starting substances and the reaction products [8]. The position of equilibrium can naturally shift markedly from one side to the other and is influenced by a change in the concentration of a reacting agent, by removal of a reaction product, and by the choice of temperature.

In the reaction of two gaseous substances the conversion can be completed by increasing the pressure when the total number of molecules decreases due to the substance transformation. Low pressure would result in a worse conversion.

Similarly to pressure changes in gas reactions, changes in concentration also result in liquids. Dilution supports reactions which are accompanied by an increase in the number of particles. Catalysts on the other hand – as already described – cannot change the position of equilibrium, but can only influence the rate of its onset.

Law of mass action

Chemical equilibrium is described by the law of mass action [8]. The equilibrium constant K is obtained from the product of the concentrations c of substances C and D, which take part in the back reaction, divided by the product of the concentrations c of substances A and B, which are involved in the direct reaction. The law of mass action applies equally to dissolving processes, chemical reactions and changes of state:



$$K = \frac{c(C) \cdot c(D)}{c(A) \cdot c(B)}.$$

When a salt is dissolved, this creates positively and negatively charged ions, i.e. cations and anions, which are in solubility equilibrium with the undissolved sediment. In a saturated solution the number of ions which dissolve equals the number of ions which precipitate again. x describes the number of cations A and y the number of anions B in the salt A_xB_y . The dissolving process sees the creation of x dissolved cations A, whose positive charge is determined by the number y of anions B. This is accompanied by the creation of y dissolved anions B with a negative charge, which is dependent on the number x of cations A.



$$K = \frac{c(A^{y+})^x \cdot c(B^{x-})^y}{c(A_xB_y)}.$$

However, the law of mass action applies in this form only to the limiting case of an ideal solution without interaction between the dissolved particles. This is approximately the case with highly diluted solutions as are created when hardly soluble salts are dissolved.

Generally speaking, the law of mass action can only be applied to solutions when the concentrations c of ions are corrected by a factor f downwards. This factor, which takes into account the mutual influence of the dissolved particles, is concentration-dependent and is known as the activity coefficient f . Thus: $f \leq 1$. The activity coefficient takes into account the tendency of ions to associate in the solution due to the difference in charge. The concentration c corrected by the activity coefficient f is called the activity a . Thus:

$$a = f \cdot c,$$

$$K = \frac{a(A^{y+})^x \cdot a(B^{x-})^y}{a(A_x B_y)}.$$

No concentration can be specified for the undissolved portion. This portion is considered to be constant and set at one. In this simplified form the law of mass action is called the solubility product K_L :

$$K_L = a(A^{y+})^x \cdot a(B^{x-})^y.$$

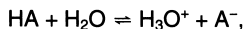
The unit of the solubility product depends on the number of different particle types which are created during the dissolving process. The lower the solubility product, the harder the salt is to dissolve. For improved clarity, as a rule the decimal logarithm of the numerical value of K_L multiplied by -1 is specified instead of the solubility product and denoted as pK_L . The greater the pK_L value, the lesser the solubility.

The proportion of salt that dissolves decays into cations and anions. The process of splitting into separate ions of different charge is frequently called dissociation. The dissolving of salts creates many charge carriers which help to conduct electric current. In aqueous solutions chemists refer to strong electrolytes.

In daily laboratory practice calculations are made approximately with concentrations c instead of with activities a . The term "concentration" is used instead of "activity" in the following text, even if the activity is retained as the physical variable in the formulae.

Acids in water

Even acids (HA, A stands for acid) dissociate in water with the delivery of hydrogen ions H^+ to the water molecules and creation of hydroxonium ions H_3O^+ (protonolysis). Here too the law of mass action can be applied [8].



$$K = \frac{a(H_3O^+) \cdot a(A^-)}{a(HA) \cdot a(H_2O)},$$

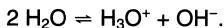
$$K_a = K \cdot a(H_2O) = \frac{a(H_3O^+) \cdot a(A^-)}{a(HA)}.$$

The concentration of water with 55.3 mol/l remains virtually constant and is incorporated into the equilibrium constant, which is then called the acidity constant or K_a value and has the unit of a concentration in mol/l. When the K_a value is high, the equilibrium moves markedly to the right. The concentration of hydroxonium ions (H_3O^+) is high. The acid therefore has a high acidic strength.

The K_a value is as a rule converted into the pK_a value, whereby the K_a value is initially divided by the standard concentration of 1 mol/l. The decimal logarithm multiplied by -1 then produces the pK_a value. Strong acids have a low or even negative pK_a value.

Acidic strength of water

Even in chemically pure water the water molecules are present in equilibrium with hydroxonium ions (H_3O^+) and hydroxide ions (OH^-).



The ionic product K_w of water is always $10^{-14} \text{ mol}^2/\text{l}^2$. Autodissociation of water is therefore very low, in which hydroxonium ions (H_3O^+) and hydroxide ions (OH^-) are created in equally high concentration of 10^{-7} Mol/l .

Disregarding the water lost through dissociation, the acidic strength K_a of water ($a(\text{H}_2\text{O}) = 55.3 \text{ mol/l}$) is obtained at:

$$K = \frac{a(\text{H}_3\text{O}^+) \cdot a(\text{OH}^-)}{a(\text{H}_2\text{O}) \cdot a(\text{H}_2\text{O})},$$

$$K_a = K \cdot a(\text{H}_2\text{O}) = \frac{a(\text{H}_3\text{O}^+) \cdot a(\text{OH}^-)}{a(\text{H}_2\text{O})},$$

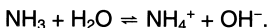
$$-\log K_a =$$

$$-\log a(\text{H}_3\text{O}^+) - \log a(\text{OH}^-) + \log a(\text{H}_2\text{O}),$$

$$\text{p}K_a = 7 + 7 + \log 55.3 = 14 + 1.74 = 15.74.$$

Bases in water

Bases on the other hand, which dissolve in water, absorb hydrogen ions (H^+) [8]. When ammonia (NH_3) is dissolved in water, ammonia molecules adopt hydrogen ions (H^+) from the water molecules. Ammonium ions (NH_4^+) and hydroxide ions (OH^-) are created:



According to the acidic constant K_a as the measure of acidic strength, it is also possible to describe the strength of bases by way of the base constant K_b . The basicity or base strength is derived from the law of mass action along similar lines to the acidic strength or more simply from the ionic product K_w of water via the relation

$$K_w = K_a \cdot K_b = 10^{-14} \text{ mol}^2/\text{l}^2 \quad \text{or}$$

$$\text{p}K_w = \text{p}K_a + \text{p}K_b = 14.$$

Corresponding acid-base pairs

When an acid, e.g. hydrochloric acid (HCl), is dissociated, the hydrochloric acid represents the acid and the acid anion in equilibrium with the acid after hydrogen-ion delivery, in this example the chloride anion (Cl^-), the base. Together they are called a corresponding acid-base pair. Similarly, a base, e.g. ammonia (NH_3), in water, with its corresponding acid, the ammonium ion (NH_4^+), forms an acid-base pair.

The $\text{p}K_a$ and $\text{p}K_b$ values are experimentally determined for many acid-base pairs (Table 3). In the event of autodissociation of water, water acts as both an acid and a base. The property of water to act as both an acid and a base is known as amphoteric behavior. When water ($\text{p}K_a = 15.74$) is the acid, after hydrogen-ion delivery hydroxide ions (OH^- , $\text{p}K_b = -1.74$) are the corresponding base. When water ($\text{p}K_b = 15.74$) acts as a base, hydroxonium ions (H_3O^+ , $\text{p}K_a = -1.74$) are the associated acid.

pH value

The pH value ("potentia hydrogenii") is defined as the negative decimal logarithm of the activity a^* of the hydroxonium ions (H_3O^+). To obtain a dimensionless quantity, the activity a of the H_3O^+ ions is divided before logarithmization by the standard concentration of 1 mol/l . The decimal logarithm then multiplied by -1 produces the pH value.

$$a^*(\text{H}_3\text{O}^+) = a(\text{H}_3\text{O}^+) [\text{mol/l}] \cdot \frac{1}{[\text{mol/l}]},$$

$$\text{pH} = -\log a^*(\text{H}_3\text{O}^+).$$

For simplification purposes the activity of the hydroxonium ions is often equated with the concentration.

The pH value is therefore a measure of the concentration of an acid; acid concentration (pH value) and acidic strength (pK_a value) do not necessarily run parallel. A diluted hydrochloric-acid solution with a concentration of 10^{-4} mol/l can despite the high acidic strength ($pK_a = -6$) be more weakly acidic ($pH = 4$) than an acetic-acid solution ($pK_a = 4.75$) with a higher concentration of 10^{-1} mol/l ($pH = 2.87$) (Figure 10).

In the event of autodissociation of water, as already described above, hydroxonium ions (H_3O^+) and hydroxide ions (OH^-) in equal concentration of 10^{-7} mol/l are created. The pH value is therefore 7, pure water is neutral.

In the laboratory in actual fact pure water often reacts slightly acidly because CO_2 from the air dissolves with the formation of carbonic acid (H_2CO_3) in the water. Carbonic acid dissociates slightly to hydroxonium ions (H_3O^+) and hydrogen-carbonate ions (HCO_3^-), pushing the pH value slightly in the acid direction.

The ratio of hydroxonium ions (H_3O^+) to hydroxide ions (OH^-) determines whether a solution is acidic, neutral or basic (Figure 10):

Acidic $a(H_3O^+) > a(OH^-)$: $pH < 7$.

Neutral $a(H_3O^+) = a(OH^-)$: $pH = 7$.

Basic $a(H_3O^+) < a(OH^-)$: $pH > 7$.

Figure 10: pH scale with examples of strong and weak acids and bases

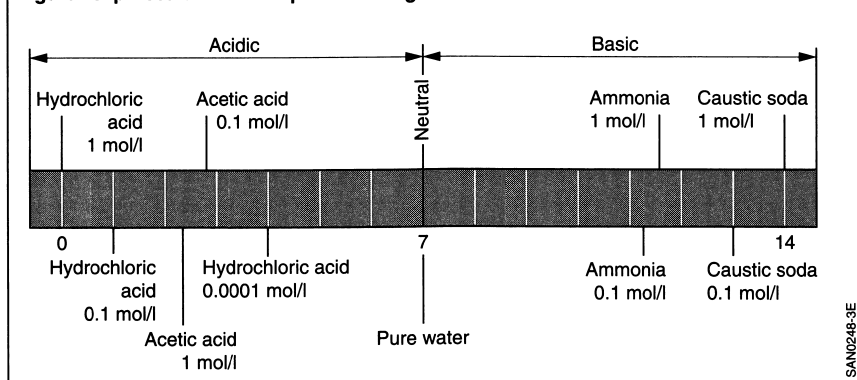


Table 3: pK_a and pK_b values of strong and weak acids

Name	Acidic strength pK_a	Corresponding acid-base pairs		Base strength pK_b
Hydrochloric acid	-6	HCl	Cl^-	20
Sulfuric acid	-3	H_2SO_4	HSO_4^-	17
Hydroxonium ion	-1.74	H_3O^+	H_2O	15.74
Nitric acid	-1.32	HNO_3	NO_3^-	15.32
Phosphoric acid	1.96	H_3PO_4	$H_2PO_4^-$	12.04
Hydrogen fluoride	3.14	HF	F^-	10.68
Acetic acid	4.75	CH_3COOH	CH_3COO^-	9.25
Carbonic acid	6.52	H_2CO_3	HCO_3^-	7.48
Ammonium ion	9.25	NH_4^+	NH_3	4.75
Hydrogen-carbonate ion	10.40	HCO_3^-	CO_3^{2-}	3.60
Hydrogen-phosphate ion	12.36	HPO_4^{2-}	PO_4^{3-}	1.64
Water	15.74	H_2O	OH^-	-1.74

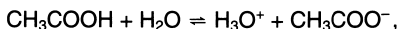
pH value, strong acids

Strong acids like hydrogen chloride dissociate completely in water, during which they transfer positively charged hydrogen ions (H^+) to water molecules with the formation of hydroxonium ions (H_3O^+). 0.1 mol (3.65 g) hydrochloric acid in 1 liter of water produces 0.1 mol hydroxonium ions and 0.1 mol chloride anions (Cl^-). A 10^{-1} molar hydrochloric-acid solution therefore has a pH value of 1.

pH value, weak acids

In the case of weak acids, a certain amount of the acid remains undissociated in the water. But because the pH value of weak acids is only determined by the share of dissociated acid molecules, the calculation of the pH value is more complicated. This will be explained here using an acetic-acid solution by way of example:

In solutions of acetic acid in water less than 1 % of all the acetic-acid molecules dissolved in water (CH_3COOH) are actually dissociated to hydroxonium ions (H_3O^+) and acetate ions (CH_3COO^-). In an aqueous acetic-acid solution, in contrast to an aqueous table-salt or hydrochloric-acid solution, the electrolyte is weak. The dissociation equilibrium can again be described using the law of mass action:



$$K_a = \frac{a(\text{H}_3\text{O}^+) \cdot a(\text{CH}_3\text{COO}^-)}{a(\text{CH}_3\text{COOH}) \cdot a(\text{H}_2\text{O})}$$

The concentration of water can again be set at one on account of the large surplus.

Acetic acid has an acidic strength K_a of $1.8 \cdot 10^{-5}$ mol/l or put another way a $\text{p}K_a$ value of 4.75. During dissociation as many hydroxonium ions (H_3O^+) as acetate ions (CH_3COO^-) are created. As the dissociation is minimal, the dissociated part compared with the total concentration of acetic acid CH_3COOH can be disregarded.

With $a(\text{H}_2\text{O}) = 1$ and $a(\text{H}_3\text{O}^+) = a(\text{CH}_3\text{COO}^-)$ it follows that:

$$K_a = \frac{a(\text{H}_3\text{O}^+)^2}{a(\text{CH}_3\text{COOH})}$$

$$a(\text{H}_3\text{O}^+)^2 = K_a \cdot a(\text{CH}_3\text{COOH})$$

$$a(\text{H}_3\text{O}^+) = \sqrt{K_a \cdot a(\text{CH}_3\text{COOH})}$$

$$-\log a(\text{H}_3\text{O}^+) =$$

$$-\frac{1}{2} (\log K_a + \log a(\text{CH}_3\text{COOH})) =$$

$$\frac{1}{2} (-\log K_a - \log a(\text{CH}_3\text{COOH})).$$

Taking into account the definitions for the pH and $\text{p}K_a$ values, the following is obtained:

$$\text{pH} = \frac{1}{2} (\text{p}K_a - \log a(\text{CH}_3\text{COOH})).$$

An acetic-acid solution with a concentration of 0.1 mol/l therefore has a pH value of

$$\begin{aligned} \text{pH} &= \frac{1}{2} (4.75 - \log 0.1) \\ &= \frac{1}{2} (4.75 + 1) = 2.87. \end{aligned}$$

pH values of other weak acids and bases can be similarly calculated.

References

- [1] K.-H. Lautenschläger, W. Weber: Taschenbuch der Chemie. Edition Harri Deutsch, 21st Edition, 2013.
- [2] M. Borlein: Kerntechnik – Grundlagen. Vogel Buchverlag, 2nd Edition, 2011.
- [3] G. Schwedt: Analytische Chemie. Wiley-VCH Verlag, 2nd Edition, 2008.
- [4] L. Spieß, G. Teichert, R. Schwarzer, H. Behnken, C. Genzel: Moderne Röntgenbeugung – Röntgendiffraktometrie für Materialwissenschaftler, Physiker und Chemiker. Springer Verlag, 3rd Edition, 2013.
- [5] K. Schwister: Taschenbuch der Chemie. Carl Hanser Verlag, 4th Edition, 2010.
- [6] M. Hesse, H. Meier, B. Zeeh: Spektroskopische Methoden in der Organischen Chemie. Thieme Verlag Stuttgart, 8th Edition, 2011.
- [7] P.W. Atkins, L. Jones: Chemie – einfach alles. Wiley-VCH Verlag, 2nd Edition, 2006.
- [8] R. Pfestorf: Chemie – Ein Lehrbuch für Fachhochschulen. Edition Harri Deutsch, 9th Edition, 2013.

Electrochemistry

Electrolytic conduction and electrolysis

When a salt is dissolved in water, the constituents of the salt are present in the water in the form of ions. These ions are charged in such a way that they move when an electric field is applied and thereby generate an electric current. This phenomenon is known as electrolytic conduction. In comparison, the current in an electric conductor such as copper or iron is transported by electrons. In addition to aqueous solutions, molten salts but also certain solids (such as zirconium oxide in the catalytic converter of a passenger car) can function as electrolytes ([1], [2] and [3]).

Negatively charged ions move towards the anode and are therefore called anions, while positively charged ions (cations) migrate towards the cathode. A chemical reaction takes place at these electrodes whereby the ions either absorb electrons from the cathode or discharge electrons to the anode. These reactions can only occur when the cathode and anode have an electrically conductive connection with each other in order to facilitate the exchange of electrons between the two.

If a battery is used as the voltage source, i.e. discharged, the electrons flow from the anode via the external electric circuit to the cathode. Thus, for the user the anode is the negative pole and the cathode is the positive pole.

Electrochemical series of metals

The intensity with which this ion reaction takes place is expressed by the electrochemical series of metals, which is set out in Table 1. The normal potential E^0 , applicable to an ion concentration of 1 mol/l, is specified. This "normal" concentration is indicated by the superscript index 0. The voltages specified in the table refer to the normal hydrogen electrode, to which a potential of 0 V is therefore allocated. The following therefore applies:

$$E^0(2\text{H}^+ + 2\text{e}^- \leftrightarrow \text{H}_2) = 0 \text{ V}.$$

In Table 1, a positive sign with E^0 indicates an electron absorption (reduction), and a negative sign an electron discharge (oxidation). For example, lithium by preference discharges an electron ($E^0 < 0$) and is oxidized to become the singly charged lithium ion Li^+ , while fluorine absorbs electrons ($E^0 > 0$) and is thus reduced.

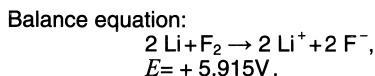
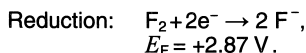
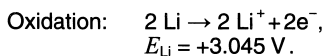


Table 1: Electrochemical series of metals ([3]) with the associated ion reactions (normal potentials at 25 °C)

Half reactions	$E^0 [\text{V}]$
$\text{Li}^+ + \text{e}^- \leftrightarrow \text{Li}$	-3.045
$\text{Na}^+ + \text{e}^- \leftrightarrow \text{Na}$	-2.714
$\text{Mg}^{2+} + 2\text{e}^- \leftrightarrow \text{Mg}$	-2.363
$\text{Al}^{3+} + 3\text{e}^- \leftrightarrow \text{Al}$	-1.662
$2\text{H}_2\text{O} + 2\text{e}^- \leftrightarrow \text{H}_2 + 2\text{OH}^-$	-0.828
$\text{Zn}^{2+} + 2\text{e}^- \leftrightarrow \text{Zn}$	-0.763
$\text{Cr}^{3+} + 3\text{e}^- \leftrightarrow \text{Cr}$	-0.744
$\text{Fe}^{2+} + 2\text{e}^- \leftrightarrow \text{Fe}$	-0.440
$\text{PbSO}_4 + 2\text{e}^- \leftrightarrow \text{Pb} + \text{SO}_4^{2-}$	-0.356
$\text{Ni}^{2+} + 2\text{e}^- \leftrightarrow \text{Ni}$	-0.250
$\text{Pb}^{2+} + 2\text{e}^- \leftrightarrow \text{Pb}$	-0.13
$\text{Sn}^{2+} + 2\text{e}^- \leftrightarrow \text{Sn}$	-0.136
$2\text{H}^+ + 2\text{e}^- \leftrightarrow \text{H}_2$	0
$\text{Cu}^{2+} + 2\text{e}^- \leftrightarrow \text{Cu}$	+0.337
$\text{Cu}^+ + \text{e}^- \leftrightarrow \text{Cu}$	+0.521
$\text{Fe}^{3+} + \text{e}^- \leftrightarrow \text{Fe}^{2+}$	+0.771
$\text{Ag}^+ + \text{e}^- \leftrightarrow \text{Ag}$	+0.799
$\text{Pt}^{2+} + 2\text{e}^- \leftrightarrow \text{Pt}$	+1.118
$4\text{H}^+ + \text{O}_2 + 4\text{e}^- \leftrightarrow 2\text{H}_2\text{O}$	+1.229
$\text{Cl}_2 + 2\text{e}^- \leftrightarrow 2\text{Cl}^-$	+1.360
$\text{Au}^{3+} + 3\text{e}^- \leftrightarrow \text{Au}$	+1.498
$\text{PbO}_2 + 4\text{H}^+ + 2\text{e}^- \leftrightarrow \text{Pb}^{2+} + 2\text{H}_2\text{O}$	+1.685
$\text{F}_2 + 2\text{e}^- \leftrightarrow 2\text{F}^-$	+2.87

An electrochemical reaction therefore always comprises the two partial steps of oxidation and reduction. Because the lithium discharges the electrons, the electrons appear on the right side of the reaction equation compared with the indication in the electrochemical series of metals (Table 1). The sign preceding the voltage must therefore be reversed for this oxidation equation. The total voltage E of the redox reaction is thus added up from the individual values. The balance equation thus does not feature any electrons, as these are only exchanged between the reacting agents.

An electrochemical reaction can thus only take place when the total voltage is positive. An acid (i.e. H^+ ions) therefore cannot dissolve copper, silver, platinum, and gold. These are referred to as noble metals. In contrast, base metals such as sodium, elemental iron, nickel and lead are attacked by acids and the metals are dissolved as ions.

To enable an electrochemical reaction to take place, at least two different reacting agents are required. The resulting electrochemical total voltage E is dependent on the concentration of ions.

Nernst equation

The Nernst equation reads:

$$E = E^0 + \frac{RT}{nF} \ln \frac{[Ox]}{[Red]}$$

$$= E^0 + \frac{0.0592 \text{ V}}{n} \log_{10} \frac{[Ox]}{[Red]} \quad (\text{at } 25^\circ \text{C}).$$

Here, [Ox] denotes the concentration of oxidized ions, [Red] the concentration of reduced ions, n the number of electrons in the reaction equation, R the gas constant, T the absolute temperature, and F the Faraday constant.

In this way, the potential can be calculated for each chemical partial reaction of the individual ion types. The total voltage of the electrolytic reactions is obtained from the sum total of all the potentials of the partial reactions. Over and above that E is dependent on the temperature.

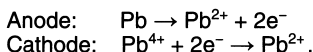
Applications

Lead battery

A motor vehicle needs a starter battery to start an engine. This starter battery is sometimes called a storage battery in that it is repeatedly discharged and charged. Regular batteries, on the other hand, are discharged only once and cannot be recharged.

Charging and discharging processes

The following electrochemical discharging reactions take place in the lead batteries used in passenger cars:



These reactions take place in sulfuric acid, where at the cathode lead oxide (PbO_2 with Pb^{4+}) is reduced to lead sulfate (PbSO_4 with Pb^{2+}), while at the anode elemental lead (Pb) is likewise oxidized to lead sulfate. The electrolyte is therefore depleted of SO_4 sulfate ions and the acid density decreases. During the charging process, the lead sulfate is converted back into lead and lead oxide.

The sulfuric acid H_2SO_4 was omitted from the above reaction equations in order to emphasize the central reactions. These reactions deliver a voltage of approx. 2.0 V (see Table 1). These reactions take place in the reverse direction when the battery is being charged while the car is running.

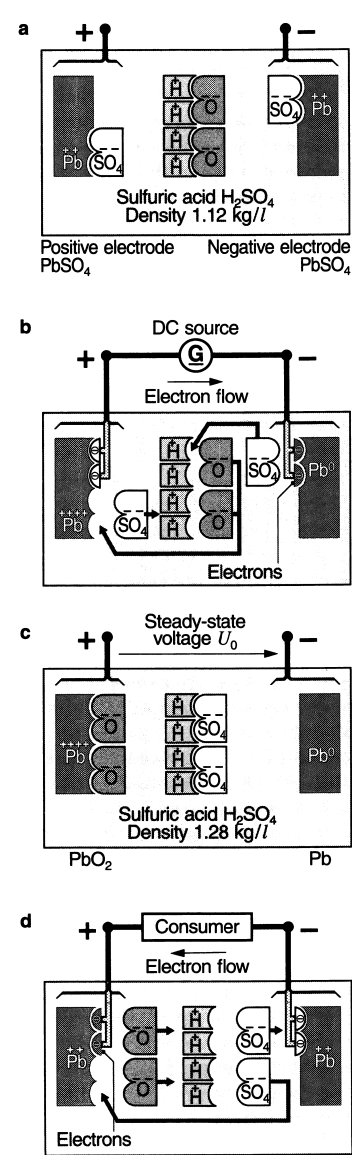
Because the electrochemical reactions described above take place very slowly, the electrodes are designed as lattices so that they have a large reaction surface. The lead and lead oxide are technically prepared as porous, sponge-like pastes and applied to these lattices.

A battery should be recharged without the voltage being allowed to drop in the process when it is fully charged. It should also be able to store a large amount of energy with minimal weight (capacity) and supply a high level of current.

Specifically the processes shown in Figure 1 occur.

Figure 1: Charging and discharging processes in a lead battery

- a) Discharged cell before charging,
- b) Charging process,
- c) Charged cell,
- d) Discharging process.



Discharged cell before charging

Located on both electrodes is PbSO_4 , which is made up of the ions Pb^{2+} and SO_4^{2-} (Figure 1a). The electrolyte has a lower density, as it is depleted of sulfate ions by the current consumption.

Charging process

Pb^{2+} is converted at the positive electrode into Pb^{4+} due to electron donation (Figure 1b). This combines with oxygen to form PbO_2 . On the other hand, elemental lead is formed at the negative electrode. Both these reactions involve the release of sulfate ions SO_4^{2-} , which with H^+ ions form sulfuric acid again and thereby increase the acid density.

The specific gravity of electrolyte can be used to indicate the state of charge of the battery (see Table 2). The accuracy of this relationship depends on battery design, electrolyte stratification, and battery wear with a certain degree of irreversible sulfating or a high degree of shedding of plate material. The electrolyte densities specified in Table 2 apply at a temperature of 20 °C; the electrolyte density drops by approximately 0.01 kg/l for every 14 K that the temperature rises and vice-versa when the temperature drops. The low value specified in Table 2 applies to high electrolyte utilization, the high value to low electrolyte utilization.

Charged cell

PbSO_4 on the positive electrode is converted into PbO_2 and PbSO_4 on the negative electrode is converted into Pb (Figure 1c). There is no further increase in electrolyte density.

If the charge voltage continues to be applied after the cell has reached a state of full charge, only the electrolytic decomposition of water occurs. This produces

Table 2: Electrolyte values of the dilute sulfuric acid in a typical auto starter battery at 20 °C

State of charge	Electrolyte density in kg/l	Freezing threshold in °C
Charged	1.28	-68
Semi-charged	1.16...1.20	-17...-27
Discharged	1.04...1.12	-13...-11

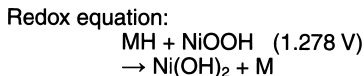
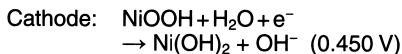
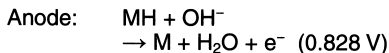
oxyhydrogen gas (oxygen at the positive electrode, hydrogen at the negative electrode).

Discharging process

The direction of current flow and the electrochemical processes during discharging are reversed in relation to charging, which results in the Pb^{2+} and SO_4^{2-} ions on both electrodes being combined to form the discharge product PbSO_4 (Figure 1d).

Nickel-metal hydride battery

In an NiMH cell (nickel-metal hydride), a metal electrode is used which can accumulate hydrogen [3]. Atomic hydrogen is created during the battery-charging process. This hydrogen is absorbed by the metal electrode (M) and a metal hydride (MH) is created. During discharging the accumulated hydrogen is oxidized on the electrode into water (Figure 2). The following reactions take place during the discharging process; these are reversed during charging:



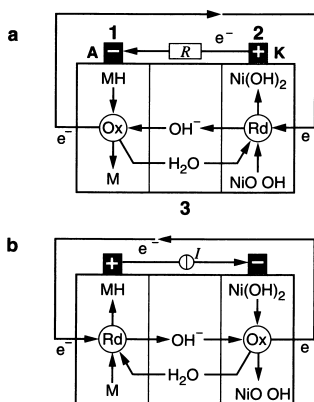
Nickel-metal hydride batteries have a high self-discharge rate of up to 1 % at room temperature, limiting their use to equipment with a short service life. They are often used in electric cars and in hybrid vehicles, since they deliver high currents and a high charging capacity while being low in weight.

Lithium-ion battery

The lithium-ion battery utilizes the reversible inclusion and withdrawal of lithium ions (Li^+) in a lattice (intercalation electrode, [3]). For example, graphite (C) in different modifications is used as anode materials and lithium metal oxides (e.g. LiCoO_2 , LiMn_2O_4) are used as the cathode (Figure 3).

Figure 2: Electrochemical processes in a nickel-metal hydride battery

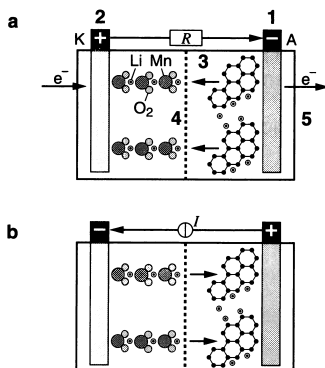
- a) Discharging,
b) Charging.
1 Anode, 2 Cathode, 3 Electrolyte.
R Resistance, *I* Charging current.



SME0686-2Y

Figure 3: Electrochemical processes in a lithium-ion battery

- a) Discharging,
b) Charging.
1 Anode, 2 Cathode, 3 Electrolyte,
4 Separator, 5 Cell housing.

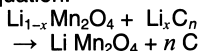


SME0687-2Y

Anode: $\text{Li}_x\text{C}_n \rightarrow n\text{C} + x\text{Li}^+ + x\text{e}^-$

Cathode: $\text{Li}_{1-x}\text{Mn}_2\text{O}_4 + x\text{Li}^+ + x\text{e}^- \rightarrow \text{LiMn}_2\text{O}_4$

Balance equation:



Depending on the anode used (Coke Carbons), which is expressed by the index n , the proportion x of the intercalated lithium changes. The electrolyte consists of waterless solvent mixtures (e.g. polyvinylidene fluoride, PVDF) and conducting salts (e.g. LiPF_6).

Lithium-ion batteries are characterized by high energy density, thermal stability and voltages up to 4.2 V.

Corrosion

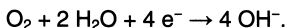
In the case of corrosion, undesirable electrochemical reactions take place in an aqueous or gaseous environment which can damage components (e.g. the body) or even result in the component failing. In this situation, for example, iron atoms in the body dissolve as ions, as a result of which the thickness of the sheet metal reduces, holes may be created in the sheet metal and, as well as the visual impairment (corrosion spots), above all the mechanical stability of the vehicle may be compromised.

The aim of the automotive industry is to protect the vehicle's iron and steel components against corrosion by means of different measures. To this end, the metal can be coated with a baser material such as zinc, which if necessary dissolves corrosively in place of the steel. Alternatively, the oxygen reaction can be stemmed by inhibitors, whereby a paint slows down the diffusion of the oxygen on the steel surface.

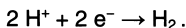
Corrosive attack

Like electrochemical reactions, a corrosive attack consists of two partial reactions – an anodic oxidation and a cathodic reduction. In the case of oxidation, the metal atom discharges electrons and dissolves as an ion or forms a deposit on the surface ("tarnishing colors") with other reacting agents. This process is the central corrosion reaction. In contrast to electrochemical reactions a corrosive attack takes place by itself, i.e. it does not need a voltage source.

In the reduction process, oxygen and hydrogen ions are often the reacting agents for anodic metal dissolution. In neutral or alkaline media, oxygen is reduced to hydroxide ions:



In acidic media, the hydrogen ions of the acids react to form hydrogen, which escapes as a gas.



The electrochemical series of metals (Table 1) provides an indication of how great the risk of corrosion is for the different metals and which materials can protect against corrosion. This table does not provide any information about the speed of the corrosion reactions. Nor does the table take into account the fact that many metals (such as aluminum, for example) in air form an oxide layer which provides temporary protection against corrosion (see also Corrosion and corrosion protection).

Oxygen concentration sensor

The oxygen concentration sensor (λ sensor) measures the residual-oxygen content in the exhaust gas and thereby regulates the air supply in order to set an optimum air/fuel mixture for combustion in the gasoline engine ($\lambda = 1$, stoichiometric combustion). To this end, a zirconium-oxide ceramic (ZrO) is used as a solid electrolyte (Figure 4), as this conducts oxygen ions at temperatures over 300 °C. Chemically inert platinum electrodes, which are applied as a porous thick layer, are used as electrodes. The voltage generated at the electrodes U_λ can be calculated with the Nernst equation. If here [Ox] and [Red] are replaced by the oxygen partial pressures $p(\text{O}_2)$ in the reference area (ambient air) and in the exhaust-gas area, the following relationship is obtained:

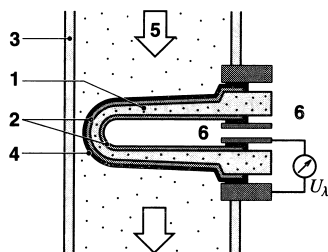
$$U_\lambda = \frac{RT}{4F} \ln \frac{p_R(\text{O}_2)}{p_A(\text{O}_2)}.$$

Here $p_A(\text{O}_2)$ denotes the oxygen partial pressure in the exhaust gas, $p_R(\text{O}_2)$ the oxygen partial pressure in the reference area, R the gas constant, T the absolute temperature, and F the Faraday constant.

In the vicinity of $\lambda = 1$, the voltage change is very great, making it possible to successfully regulate to this value.

Figure 4: λ sensor (finger-type sensor) in the exhaust pipe

- 1 Zirconium-oxide ceramic, 2 Electrodes,
- 3 Exhaust pipe,
- 4 Porous ceramic protective layer,
- 5 Exhaust gas, 6 Outside air (reference air).



UNIK2115Y

References

- [1] E. Fluck, R.C. Brasted: *Allgemeine und Anorganische Chemie*. 6th Ed., UTB No. 53, Quelle & Meyer, Heidelberg, 1987.
- [2] C.E. Mortimer, U. Müller: *Chemie*. 12th Ed., Thieme, Stuttgart, 2015.
- [3] C.H. Hamann, W. Vielstrich: *Elektrochemie*. 4th Ed., Wiley-VCH, 2005.

Mathematics

Numbers

Quantities

Numbers are divided into natural numbers $\mathbb{N} = \{0, 1, 2, 3, \text{etc.}\}$, whole numbers $\mathbb{Z} = \{\dots, -3, -2, -1, 0, +1, +2, +3, \dots\}$, rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$.

In addition to all whole numbers, rational numbers also include all fractions whose numerators and denominators are whole numbers. In addition to the rational numbers, real numbers also include all (infinitely many) numbers between fractions. Two examples of real numbers are the circular constant pi ($\pi = 3.14159\dots$) and the Eulerian number e ($e = 2.718281\dots$). Complex numbers are an extension of real numbers. These are explained in detail below (for more information, refer to [1], [2] and [3]).

Number systems

Physical and technical quantities are described by their numerical value and their unit. Numbers are usually represented in the decimal system, i.e. using base 10. Other common number systems are the binary and the hexadecimal systems, which are based on the numbers 2 and 16, respectively. While the decimal system features the digits 0, 1, 2 and onwards through 9, only the digits 0 and 1 exist in the binary system. The hexadecimal system uses, in addition to the digits 0 through 9, the letters A through F to represent the numbers 10 through 15 (for conversion, see Tables 1 and 2).

The binary and hexadecimal systems are used primarily in the field of information technology (IT), since a computer can only process the two states “power off” (“0”) and “power on” (“1”). These two states form the basis of the binary system. When eight binary digits are combined into one byte, it is then possible to represent the numbers 0 to 255, which corresponds to the hexadecimal digits 0 – FF.

Functions

The following section describes elementary mathematics functions. Their most important properties, such as their domain and value set, their behavior for very large and very small x values, their zero points, their derivations and the fundamental arithmetic operations with them are depicted.

These functions have been selected from a very large number of functions since they can be used to illustrate many technical processes, such as the geometric correlations between a connecting rod and moving mechanical parts (see Internal-combustion engines, Crankshaft gears), vibrations and oscillations in the vehicle (see Undercarriage, Basic principles of vibration characteristics), distance measurements of the vehicle relative to other cars and people (see Driver-assistance systems, Parking systems).

Table 1: Decimal and binary system

Decimal	Binary
0	0
1	1
2	10
3	11
4	100
8	1000
9	1001
15	1111
16	10000
32	100000
64	1000000
255	11111111

Table 2: Decimal and hexadecimal system

Decimal	Hexadecimal
0, 1–9	0, 1–9
10, 11–15	A, B–F
16	10
17	11
30	1E
31	1F
32	20
255	FF
4096	1000
65535	FFFF

Polynomial

A polynomial of the n th degree consists of $(n+1)$ summands with real (or complex) coefficients $a_0, a_1, \dots, a_n$ and the associated monomials x^i :

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

where $a_i \in \mathbb{R}$ (or $a_i \in \mathbb{C}$)
and the domain range $D_f = \mathbb{R}$ (or $D_f = \mathbb{C}$)
and the value set $W_f = \mathbb{R}$.

A polynomial of the n th degree may have up to n zero points and $n-1$ local extrema.

Straight line

A straight line is a polynomial of the first degree ($n=1$):

$$\begin{aligned} f(x) &= a_0 + a_1x \quad \text{or} \\ y &= mx + t \quad (\text{with slope } m \text{ and} \\ &\quad \text{axis intercept } t). \end{aligned}$$

The zero point lies at $x_0 = -\frac{t}{m}$ (for $m \neq 0$).

Parabola

A parabola is a polynomial of the second degree ($n=2$, quadratic function):

$$\begin{aligned} f(x) &= a_0 + a_1x + a_2x^2 \quad \text{or} \\ y &= ax^2 + bx + c. \end{aligned}$$

$$\text{Zero points: } x_{1/2} = \frac{1}{2a} (-b \pm \sqrt{b^2 - 4ac})$$

Table 3: Polynomial

Definition range D_f , value set W_f $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ $D_f = \mathbb{R}, W_f = \mathbb{R}$ (or subset).
Figure 1: Straight line (dashed) and parabola as polynomials of the first and second degree

A parabola may have zero, one or two zero points.

In Table 3 and Figure 1 shows by way of example a straight line (polynomial of the first degree) and a parabola (polynomial of the second degree).

Root function

The root function $f(x) = \sqrt{x} = x^{1/2}$ (Table 4 and Figure 2) is the inverse function of the quadratic function $f(x) = x^2$. It is needed for tasks such as solving quadratic equations (e.g. searching for the zero point of a second-degree polynomial, see Polynomial).

This also results in the root term, which is used in the calculating the resonant frequency and the damping ratio of an oscillation (see Oscillations).

Similarly, for each monomial $f(x) = x^n$, $n \in \mathbb{N}$, there is a strongly monotonically growing inverse function

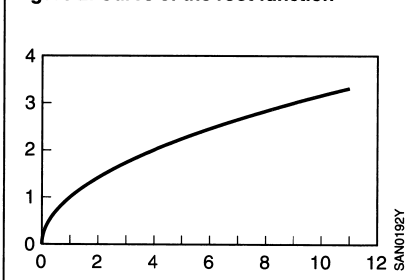
$$f^{-1}(x) = \sqrt[n]{x} = x^{1/n},$$

which is only defined for $x \geq 0$.

Table 4: Root function

Definition range D_f , value set W_f and behavior. $f(x) = \sqrt{x} = x^{1/2}$ $D_f = \mathbb{R}_0^+$ $W_f = \mathbb{R}_0^+$ $x \rightarrow +\infty: f(x) \rightarrow +\infty$	Properties. $\sqrt{x} \cdot \sqrt{y} = \sqrt{xy}$ $\sqrt{x^n} = (\sqrt{x})^n, n \in \mathbb{N}$
--	---

Figure 2: Curve of the root function



Absolute value and sign function

Each real number x can be broken down into its sign (+/−) and its absolute value $|x|$.

$x = \operatorname{sgn}(x) \cdot |x|$.

For example, $\operatorname{sgn}(-3) = -1$ and $\operatorname{sgn}(+4) = +1$. The absolute value indicates the distance of the number x from a source of 0. The absolute value function (Table 5 and Figure 3) can be represented as

$$f(x) = |x| = \begin{cases} x & \text{for } x \geq 0 \\ -x & \text{for } x < 0. \end{cases}$$

Exponential function

The exponential function is a very important function in mathematics, physics and technology, because it is the only function that is identical to its own derivative. This central property is used in solving linear differential equations such as the harmonics equation (see Oscillations).

Both the sine and cosine functions are related to the complex exponential function (see Complex numbers). This means the exponential function can be used to describe oscillations.

$f(t) = e^{-\gamma t + i \omega t} = e^{-\gamma t} (\cos \omega t + i \sin \omega t)$.

Here, the real negative exponent $-\gamma t$ represents the damping of the oscillation, while the complex exponent $i \omega t$ reflects the periodic quantity, which becomes clear from the representation of the sine and cosine functions (see Complex numbers, Linear differential equations).

Furthermore, laws of growth (e.g. interest and compound interest calculations) and laws of decay (e.g. radioactive decay) can be expressed with the exponential function. The charging and discharging processes of capacitors also follow an exponential curve (see Capacitor). The final compression pressure, final compression temperature and efficiency of a reciprocating-piston engine are exponentially related to the polytropes – more specifically, the adiabatic exponents (see Reciprocating-piston engine).

Table 6 and Figure 4 show the domain and value set, as well as the properties of the exponential function.

Table 5: Absolute value function

Definition range D_f Value set W_f and behavior.
$f(x) = \operatorname{sgn}(x)$
$D_f = \mathbb{R}$ $W_f = \mathbb{R}_0^+$ $x \rightarrow \pm \infty : f(x) \rightarrow +\infty$
Figure 3: Curve of the absolute value function

Table 6: Exponential function

Definition range D_f and behavior.	Properties.
$f(x) = e^x ; D_f = \mathbb{R}$ $x \rightarrow -\infty : f(x) \rightarrow 0$ $x \rightarrow +\infty : f(x) \rightarrow +\infty$ $f(0) = e^0 = 1$	$e^a e^b = e^{a+b}$ $(e^a)^b = e^{ab}$ $\frac{d}{dx} e^x = e^x$
Euler's number $e = 2.71828\dots$	
Figure 4: Curve of the exponential function	

Logarithm function

The logarithm function is the inverse function of the exponential function. This is needed in order to solve equations such as the following:

$$2^x = 8 \Rightarrow x = \log_2 8 \Rightarrow x = 3.$$

Table 7 and Figure 5 show the domain and value set, as well as the properties of the logarithm function.

Application examples

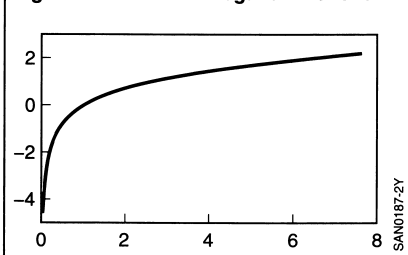
The Nernst equation provides one example of application of the logarithm function. This equation calculates the voltage value from the concentration of oxygen in the surroundings and the exhaust gas, which is necessary for λ regulation (see λ sensor).

In acoustics, decibel level (dB) is defined using the logarithm of the sound pressure (see Acoustics, Decibels). Similarly, logarithms are also used to determine the sound power level and sound intensity level.

Table 7: Logarithm function

Definition range D_f and behavior.	Properties.
$f(x) = \ln x$ $= \log_e x$; $D_f = \mathbb{R}^+$	$\ln a + \ln b = \ln ab$
$x \rightarrow 0: f(x) \rightarrow -\infty$	$\ln a - \ln b = \ln \frac{a}{b}$
$x \rightarrow +\infty: f(x) \rightarrow \infty$	$c \ln a = \ln a^c$
$f(1) = \ln 1 = 0$	$\frac{d}{dx} \ln x = \frac{1}{x}$

Figure 5: Curve of the logarithm function



Converting different logarithms

The logarithm of the number z for the base a can be converted to the base b as follows ($a, b > 0$):

$$\log_a z = \frac{\log_b z}{\log_b a}$$

This results in $\lg z = \log_{10} z$ for the common (decimal) logarithm and $\lg z = \log_2 z$ for the binary logarithm:

$$\log_a z = \frac{\ln z}{\ln a} = \frac{\lg z}{\lg a}.$$

Trigonometric functions

Angular measurement (radians)

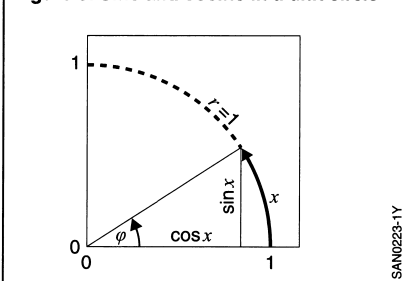
In mathematics, angles are usually specified as the measure of an arc in radians and rarely in degrees. So angle $\varphi = 360^\circ$ is equivalent to a arc of $x = 2\pi$. The circumference of a circle with a radius of 1 shares this value.

This yields the following conversions between angle φ in degrees and angle x in radians:

$$\begin{aligned} \frac{\varphi}{360^\circ} &= \frac{x}{2\pi}, \\ \Rightarrow \varphi &= \frac{180^\circ}{\pi} x, \\ \Rightarrow x &= \frac{\pi}{180^\circ} \varphi. \end{aligned}$$

The unit “rad” is often added to an arc measurement to make it clear that the information relates to an angle. Arc x of the angle φ is shown in Figure 6. Angle φ of the associated arc x is shown as a curved arrow.

Figure 6: Sine and cosine in a unit circle



Sine and cosine functions

In a right triangle, the sine of angle φ or arc x is equal to the ratio of the opposite side to the hypotenuse. The cosine is the ratio of the adjacent side to the hypotenuse.

In a right triangle with a hypotenuse with length $r = 1$ (unit circle), the adjacent side corresponds to the cosine and the opposite side corresponds to the sine (Figure 6).

Many periodic processes such as oscillations can be expressed through the sine or cosine function (Table 8 and Figure 7). If the variable x is replaced by

$$x = \frac{2\pi}{T} t \text{ and } \omega = \frac{2\pi}{T}$$

where the variable t usually stands for the time, then T is the period of the oscillation. The frequency f is the reciprocal of the period:

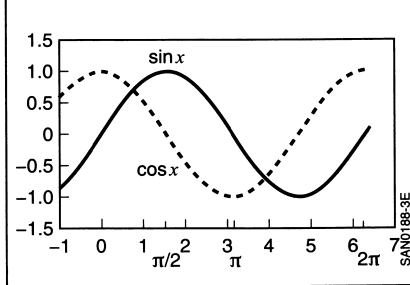
$$f = \frac{1}{T}$$

The angular frequency ω still contains the factor 2π ; thus, it specifies the angle covered (in radians) per unit of time.

Table 8: Sine and cosine functions

Definition ranges D_f, D_g , Value sets W_f, W_g and behavior.	Properties.
$f(x) = \sin x ; D_f = \mathbb{R}$	$\sin(x + 2\pi) = \sin x$
$g(x) = \cos x ; D_g = \mathbb{R}$	$\cos(x + 2\pi) = \cos x$
$W_f = W_g = [-1; +1]$	$\sin(x \pm \frac{\pi}{2}) = \pm \cos x$
Period length: 2π	$\cos(x \pm \frac{\pi}{2}) = \mp \sin x$
	$\sin^2 x + \cos^2 x = 1$

Figure 7: Curves of the sine and cosine functions



When a car drives on a street or a country road, it is constantly subjected to large and small bumps and shocks. Therefore, the undercarriage must be constructed so that it absorbs or compensates for the bumps in the road. In addition, the vibrational characteristics of the shock absorbers should be designed with precision (see Undercarriage, Basic principles of vibration characteristics).

Superpositions of oscillations are important in many applications, such as water waves, sound waves and the superposition of alternating electrical current (AC). In addition, the wave is usually formulated as a sine or cosine function

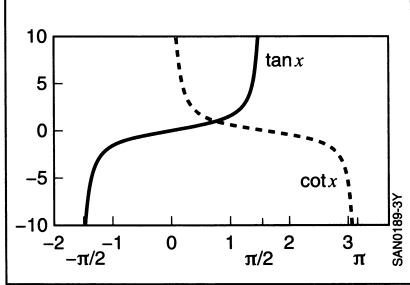
$$f(t) = A \sin(\omega t + \varphi) \text{ or } f(t) = A \cos(\omega t + \varphi)$$

with (positive) amplitude A and phase shift φ .

Table 9: Tangent and cotangent functions

Definition ranges D_f, D_g , Value sets W_f, W_g , and behavior.	Properties.
$f(x) = \tan(x) = \frac{\sin(x)}{\cos(x)}$	$\tan(x + \pi) = \tan(x)$
$g(x) = \cot(x) = \frac{\cos(x)}{\sin(x)}$	$\cot(x + \pi) = \cot(x)$
$D_f = \mathbb{R} \setminus \{\frac{\pi}{2} + k \cdot \pi, k \in \mathbb{Z}\}$	$\tan(x + \frac{\pi}{2}) = -\cot(x)$
$D_g = \mathbb{R} \setminus \{k \cdot \pi, k \in \mathbb{Z}\}$	$\cot(x + \frac{\pi}{2}) = -\tan(x)$
$W_f = W_g = \mathbb{R}$	$\cot(x) = \frac{1}{\tan(x)}$
Period length: π	

Figure 8: Curves of tangent and cotangent functions



If oscillations with the same angular frequency are superimposed, they can be amplified (through constructive interference), weakened or even completely eliminated (through destructive interference). This depends on the phase relation and the amplitudes of the relevant oscillations.

For example, destructive interference is utilized as “anti-noise” at airports. This involves having speakers produce sounds with the same frequency as the turbine noise from airplanes. The anti-noise and noise are out of phase with one another, canceling each other out, reducing the noise for nearby residents.

Tangent and cotangent functions

In a right triangle, the tangent results from the ratio of the opposite side to the adjacent side. Therefore, the following relationship can be derived:

$$\tan x = \frac{\sin x}{\cos x}$$

The cotangent results from the ratio of the adjacent side to the opposite side. This means that

$$\cot x = \frac{\cos x}{\sin x}$$

Table 9 and Figure 8 show the properties and the curves of both of these functions. The most important properties of trigonometric functions are summarized in Table 10.

Arc functions

The arc functions are the inverse of trigonometric functions. For example, the following equation can be solved using the arcsine:

$$\sin x = 0.5 \quad \text{where } x \in [0; \pi/2]$$

$$\Rightarrow x = \arcsin(0.5) = \pi/6.$$

Similarly, there are also an arccosine, arc-tangent and arccotangent.

In contrast to the sine, cosine, tangent and cotangent, which (with the exception of gaps) are defined for any $\mathbb{R}$, the arc functions may have a finite interval as a domain, in particular the value range of the original function. These properties and others are compiled in Table 11. The arc functions are not periodic.

Table 10: Properties of trigonometric functions, $k \in \mathbb{Z}$

	$\sin(x)$	$\cos(x)$	$\tan(x)$	$\cot(x)$
D_f	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R} \setminus \{x x = \frac{\pi}{2} + k\pi\}$	$\mathbb{R} \setminus \{x x = k\pi\}$
W_f	$[-1; +1]$	$[-1; +1]$	$\mathbb{R}$	$\mathbb{R}$
Period	2π	2π	π	π
Symmetry	Odd	Even	Odd	Odd
Zero points	$x_0 = k\pi$	$x_0 = \frac{\pi}{2} + k\pi$	$x_0 = k\pi$	$x_0 = \frac{\pi}{2} + k\pi$
Maxima	$x_{\max} = \frac{\pi}{2} + k \cdot 2\pi$	$x_{\max} = k \cdot 2\pi$	-	-
Minima	$x_{\min} = \frac{3}{2} + k \cdot 2\pi$	$x_{\min} = \pi + k \cdot 2\pi$	-	-
Poles	-	-	$x_{\text{Pole}} = \frac{\pi}{2} + k\pi$	$x_{\text{Pole}} = k\pi$

Table 11: Properties of arc functions

	$\arcsin(x)$	$\arccos(x)$	$\arctan(x)$	$\operatorname{arccot}(x)$
D_f	$[-1; +1]$	$[-1; +1]$	$\mathbb{R}$	$\mathbb{R}$
W_f	$[-\frac{\pi}{2}; \frac{\pi}{2}]$	$[0; \pi]$	$]-\frac{\pi}{2}; \frac{\pi}{2}[$	$]0; \pi[$
Symmetry	Odd	Point symmetry for $P(0; \frac{\pi}{2})$	Odd	Point symmetry for $P(0; \frac{\pi}{2})$
Zero points	$x_0=0$	$x_0=1$	$x_0=0$	-
Monotony	Strictly monotonically increasing	Strictly monotonically decreasing	Strictly monotonically increasing	Strictly monotonically decreasing
Asymptotes	-	-	$y=\pm \frac{\pi}{2}$	$y=0; y=\pi$

Equations in the plane triangle

Figure 9 shows a triangle with any lengths of sides a , b and c and any angles α , β and γ . The relationship between these quantities is described by the following equations.

Angular sum:
 $\alpha + \beta + \gamma = 180^\circ$.

Law of sines:
 $a : b : c = \sin \alpha : \sin \beta : \sin \gamma$.

Law of cosines:
 $c^2 = a^2 + b^2 - 2ab \cos \gamma$.

Pythagorean theorem:
 $c^2 = a^2 + b^2$, where $\gamma = 90^\circ$.

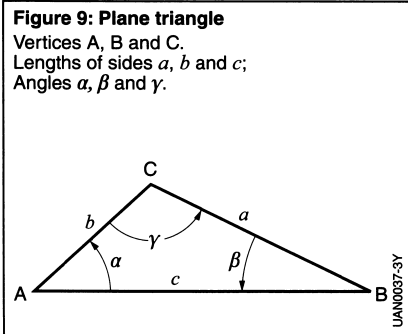
The relationships between angles in a triangle play an important role in the design of components. One example of this would be the gear design in a reciprocating piston engine, where the piston stroke and length of the connecting rod are related to one another by the angle of the crankshaft position and the leg angle of the connecting rod (see Internal-combustion engines, Gear design).

Complex numbers

In real numbers, a root can be extracted from any positive number; this is not possible with negative numbers. In order to bypass this restriction, real numbers are changed into complex numbers. Their central element is the imaginary unit i , where:

$$i^2 = -1.$$

A complex number z has a real component (Re) and an imaginary component (Im). The imaginary component is multiplied by the imaginary unit i . It is possible with the coordinates a and b to represent z in Cartesian form. Using a simple transformation, this can be converted into the polar coordinates r and φ (distance from the origin of the coordinates and the angle to the X -axis) (Figure 10).



Complex number z in Cartesian form:

$$z = a + ib \quad \text{where } i^2 = -1.$$

$$\operatorname{Re} z = a, \operatorname{Im} z = b.$$

Complex number z in polar coordinates:

$$z = r e^{i\varphi}$$

$$\text{where } r = \sqrt{a^2 + b^2}, \tan \varphi = \frac{b}{a}$$

$$a = r \cos \varphi, b = r \sin \varphi.$$

Complex exponential function

(Euler's formula)

$$e^{i\varphi} = \cos \varphi + i \sin \varphi$$

Calculation rules for complex numbers

In the case of complex numbers, the same calculation rules that are used for real numbers apply, whereby addition is simpler to carry out in Cartesian notation and multiplication is simpler in polar notation.

Addition:

$$z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2).$$

Multiplication:

$$z_1 z_2 = (r_1 r_2) e^{i(\varphi_1 + \varphi_2)}$$

In many cases, complex numbers aid in making complicated mathematical problems easier to solve. Frequently, only the real part or the imaginary part of the complex solution is necessary in finding the real solution.

For example, oscillation processes can be expressed using linear differential equations (see Differential equations), such as the oscillation and damping properties of shock absorbers or the current strength and voltage in an alternating-current circuit. A solution can be found quickly with the help of a complex exponential function. Finally, Euler's formula for exponential functions is used to convert a complex solution into a real solution.

Coordinate systems

Coordinate systems on a plane

In the previous segment, a complex number z is either represented by both of its coordinates x and y or by its absolute value $r = |z|$ and the angle φ in relation to the x -axis. This number can be graphically represented as a point on a plane (Figure 10).

Similarly, all the points on a plane can be expressed by their Cartesian coordinates (x and y) or by their polar coordinates (r and φ), and converted into one another.

Conversion of Cartesian coordinates to polar coordinates:

$$r = \sqrt{x^2 + y^2}, \quad x, y \in \mathbb{R},$$

$$\tan \varphi = \frac{y}{x}.$$

Conversion of polar coordinates to Cartesian coordinates:

$$x = r \cos \varphi, \quad r \in \mathbb{R}_0^+, \varphi \in [0; 2\pi],$$

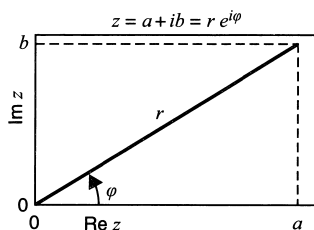
$$y = r \sin \varphi.$$

This conversion always yields a unique value, meaning that for each x - y pair, there is an exact r - φ pair. r is only $= 0$ at the coordinate origin $(0, 0)$, but the angle φ is undefined. However, in practice this does not amount to a limitation.

Cartesian coordinates can be used for straight-line movements. Naturally, circular motion can also be represented us-

Figure 10: Complex number z

Diagram in Cartesian coordinates (a and b) and polar coordinates (r and φ).



ing Cartesian coordinates. However, the corresponding equations are significantly more complicated in comparison to a description in polar coordinates (see for example Fundamentals of vehicle engineering, Translational movement).

Coordinate systems in three dimensions

Each point in a three-dimensional space has three Cartesian coordinates (x, y, z) . This representation is used frequently when the system to be described is at a right or oblique angle. However, if the problem relates to a rotationally symmetrical or a spherically symmetrical system, using cylindrical or spherical coordinates is recommended.

In Figure 11, the x , y and z coordinate axes are delineated. Here, the distance r of point P to the origin and the two angles θ and φ , which can be measured from the z or the x -axis. r , θ and φ are labeled as spherical coordinates or spatial polar coordinates.

With the exception of the origin $(0, 0, 0)$, the coordinates for each point are clearly defined in both Cartesian and polar coordinates. Conversion between both coordinate system occurs as follows.

Conversion of Cartesian coordinates to polar coordinates:

$$r = \sqrt{x^2 + y^2 + z^2}, \quad x, y, z \in \mathbb{R},$$

$$\tan \varphi = \frac{y}{x},$$

$$\cos \theta = \frac{z}{r} = \frac{z}{\sqrt{x^2 + y^2 + z^2}}.$$

Conversion of polar coordinates to Cartesian coordinates:

$$\begin{aligned} x &= r \cos \varphi \sin \theta, \\ r &\in \mathbb{R}_0^+, \varphi \in [0; 2\pi], \theta \in [0; \pi], \\ y &= r \sin \varphi \sin \theta, \\ z &= r \cos \theta. \end{aligned}$$

The surface of the Earth is broken down cartographically into geographic coordinates. The geographic length is the angle φ in spherical coordinates. The geographic width is determined by an angle of -90° to $+90^\circ$ relative to the equator. This is equivalent to the angle θ in spherical coordinates, which is measured starting from the north pole and is between 0° and 180° (0 and π in radians).

Cylindrical coordinates

If an object or a system has rotational symmetry, cylindrical coordinates are utilized (Figure 12). The x and y coordinates, just like the polar coordinates on a plane, are converted to the radius ρ ; the z component of the cylindrical coordinates

Figure 11: Spherical coordinates of a point

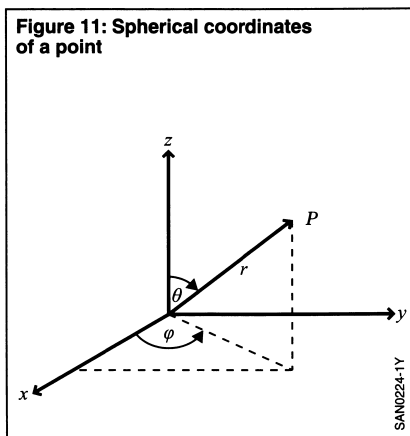
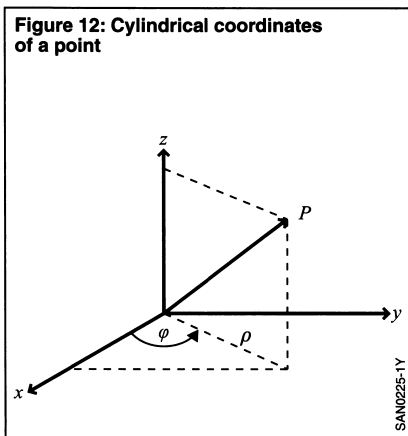


Figure 12: Cylindrical coordinates of a point



is identical to the z of the Cartesian coordinates.

Just like the spherical coordinates, any point can be clearly expressed in cylindrical coordinates, with the exception of the origin $(0, 0, 0)$.

Conversion of Cartesian coordinates to cylindrical coordinates:

$$\rho = \sqrt{x^2 + y^2}, \quad x, y, z \in \mathbb{R},$$

$$\tan \varphi = \frac{y}{x},$$

$$z = z.$$

Conversion of cylindrical coordinates to Cartesian coordinates:

$$\begin{aligned} x &= \rho \cos \varphi, \quad \rho \in \mathbb{R}_0^+, \varphi \in [0; 2\pi], z \in \mathbb{R}, \\ y &= \rho \sin \varphi, \\ z &= z. \end{aligned}$$

Just like the polar coordinates, cylindrical coordinates are calculated on a plane and they are expanded by the addition of the z -axis.

Cylindrical coordinates can be used in the design of rotationally symmetrical objects, such as tubes, engine cylinders, screws, nuts, and anti-friction and friction bearings.

Vectors

A distinction is made between scalar and vector variables for physical and technical concepts. Examples of scalars include mass, temperature and pressure. By comparison, vectors have a direction, such as velocity, power and electric field. Vectors, their most important arithmetic operations and calculation rules are shown below.

Representation of vectors

A vector in a three-dimensional space has the three components a_x , a_y and a_z for the x -, y and z direction. In mathematics, this is described as follows:

$$\vec{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}.$$

Calculation rules

When adding and multiplying vectors with a number, the same calculation rules for adding and multiplying numbers apply.

Addition

For adding two vectors with coordinates as follows

$$\vec{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \quad \text{and} \quad \vec{c} = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix}$$

results in:

$$\vec{a} + \vec{c} = \begin{pmatrix} a_x + c_x \\ a_y + c_y \\ a_z + c_z \end{pmatrix}$$

The following laws apply:

- Closure property: The sum and the difference of two vectors is also a vector.
- Commutative property:
 $\vec{a} + \vec{c} = \vec{c} + \vec{a}.$
- Associative property:
 $\vec{a} + (\vec{c} + \vec{n}) = (\vec{a} + \vec{c}) + \vec{n}.$
- For any two vectors $\vec{a}$ and $\vec{c}$ there is always one vector $\vec{z}$ so that:
 $\vec{a} + \vec{z} = \vec{c},$ in other words $\vec{z} = \vec{c} - \vec{a}.$

Multiplication of a vector with a scalar

When multiplying vector a by scalar $\lambda \in \mathbb{R}$, the following is true:

$$\lambda \vec{a} = \begin{pmatrix} \lambda a_x \\ \lambda a_y \\ \lambda a_z \end{pmatrix}$$

For vector $\vec{c}$, the following is true:

$$\vec{c} = \lambda \vec{a} \Rightarrow |\vec{c}| = |\lambda| \cdot |\vec{a}|.$$

For vectors $\vec{a}$ and $\vec{c} \in \mathbb{R}^3$ and scalars λ and $\mu \in \mathbb{R}$, the following laws apply:

- Closure property: The product of a vector and a scalar is also a vector.
- Associative property: $(\lambda \cdot \mu) \cdot \vec{a} = \lambda \cdot (\mu \cdot \vec{a})$.
- Distributive law:
 - $(\lambda + \mu) \cdot \vec{a} = \lambda \cdot \vec{a} + \mu \cdot \vec{a}$,
 - $\lambda \cdot (\vec{a} + \vec{c}) = \lambda \cdot \vec{a} + \lambda \cdot \vec{c}$.
- Multiplying by the neutral element:
 - $1 \cdot \vec{a} = \vec{a}$.

Scalar product

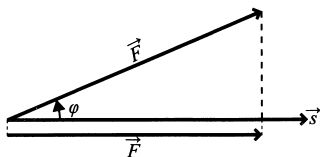
When multiplying two vectors with one another, two different products are defined: the scalar product and the vector product.

The result of the scalar product is a number (scalar). For example, physical work is defined as a scalar product of the force vector $\vec{F}$ and the path vector $\vec{s}$ (Figure 13). Only the components of $\vec{F}$ that point in the direction of $\vec{s}$ contribute to the work. This corresponds to the offset projection depicted in Figure 13. Thus, the following is true for the absolute value of the projection

$$|\vec{F}_s| = |\vec{F}| \cos \varphi$$

Figure 13: Scalar product

The scalar product of both vectors $\vec{F}$ and $\vec{s}$ with the projection $\vec{F}_s$ of the vector $\vec{F}$ on $\vec{s}$. φ is the angle between both vectors $\vec{F}$ and $\vec{s}$.



SAN0226Y

The scalar product is defined as

$$\begin{aligned} \vec{a} \cdot \vec{c} &= |\vec{a}| \cdot |\vec{c}| \cdot \cos \varphi \\ &= a_x c_x + a_y c_y + a_z c_z. \end{aligned}$$

In order to clarify the product, the multiplication point “ $\cdot$ ” is placed between both vectors.

The calculation rules for the scalar product are the same as the calculation rules for numbers.

- Commutative property:
 - $\vec{a} \cdot \vec{c} = \vec{c} \cdot \vec{a}$.
- Distributive law:
 - $\vec{a} \cdot (\vec{c} + \vec{n}) = \vec{a} \cdot \vec{c} + \vec{a} \cdot \vec{n}$.
- Multiplication with a scalar, associative property:
 - $\lambda (\vec{a} \cdot \vec{c}) = (\lambda \vec{a}) \cdot \vec{c} = \vec{a} \cdot (\lambda \vec{c})$.

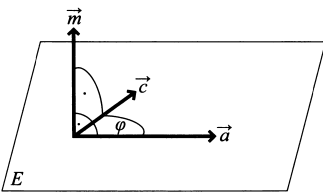
For example, the scalar product is used to calculate the absolute value (the length) of a vector. In addition, the angles between the vectors can be determined using the same method. The scalar product in particular is used in order to find out if two vectors are perpendicular to each other. In this case, the scalar product is zero.

Cross product

A new vector results from the cross product (outer product) of two vectors; this new vector is perpendicular to both output vectors (Figure 14). Its length equals the area of a parallelogram spanning both initial vectors. The vectors $\vec{a}$, $\vec{c}$ and $\vec{m}$ form what is known as a right-handed system. According to the right-hand rule, the thumb points in the direction of $\vec{a}$, the index finger

Figure 14: Cross product

Both vectors $\vec{a}$ and $\vec{c}$ span the length of a plane E . The cross product (vector $\vec{m}$) formed from $\vec{a}$ and $\vec{c}$ points perpendicular to both output vectors. φ is the angle between $\vec{a}$ and $\vec{c}$.



SAN0227Y

in the direction of $\vec{c}$ and the resulting vector $\vec{m}$ in the direction of the middle finger on the right hand.

In order to be able to distinguish the cross product from the scalar product, an "x" is placed between the vectors as a multiplication symbol.

The calculation rules for the cross product are summarized below:

$$\vec{m} = \vec{a} \times \vec{c} \quad \text{where}$$

$$|\vec{m}| = |\vec{a}| \cdot |\vec{c}| \cdot \sin \varphi = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ a_x & a_y & a_z \\ c_x & c_y & c_z \end{vmatrix} = \begin{bmatrix} a_y c_z - a_z c_y \\ a_z c_x - a_x c_z \\ a_x c_y - a_y c_x \end{bmatrix}$$

$$\vec{m} \perp \vec{a}, \vec{m} \perp \vec{c},$$

$\vec{a}, \vec{c}, \vec{m}$ are "right-handed"

with the unit vectors $\vec{e}_x, \vec{e}_y$ and $\vec{e}_z$ in x, y and z direction.

When applying the calculation rules, note that the leading sign is inverted if the order of the vectors is reversed (anticommutativity):

- Anticommutativity:
 $\vec{a} \times \vec{c} = -\vec{c} \times \vec{a}.$
- Distributive law:
 $\vec{a} \times (\vec{c} + \vec{n}) = \vec{a} \times \vec{c} + \vec{a} \times \vec{n}.$
- Multiplication with a scalar, associative property:
 $\lambda(\vec{a} \times \vec{c}) = (\lambda \vec{a}) \times \vec{c} = \vec{a} \times (\lambda \vec{c}).$

The cross product becomes zero when both vectors point in the same or opposite direction (i.e. when they are collinear). This lets you use the cross product to check if two vectors span one plane. In this case, the cross product does not equal zero.

Therefore, the cross product of a vector with itself is zero.

$$\vec{a}, \vec{c} \neq \vec{0} \quad \text{where} \quad \vec{a} \times \vec{c} = \vec{0} \Leftrightarrow \vec{a} \parallel \vec{c},$$

$$\vec{a} \times \vec{a} = \vec{0}.$$

Differential and integral calculus

Differentiation of functions

First derivative

The zero points, discontinuities and the behavior of a function can be determined at the margins of the domain using the function term $y = f(x)$. Finding out how steep a function is or where the function attains its maximum or minimum values is often of interest as well. This can be described using differentiation (the derivative) of a function.

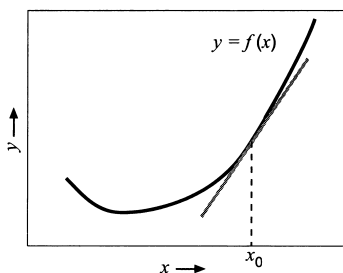
The slope (steepness) of a function $y = f(x)$ is given via the tangent of this function (Figure 15). If a function attains a maximum or a minimum value, then the tangent is horizontal at that point. In order to calculate the slope of a function $f(x)$, it must be derived (differentiated). The following two notation systems have been established for the first derivative of a function $f(x)$ to accomplish this:

$$\frac{d}{dx} f(x) = f'(x).$$

The derivatives of elementary functions are summarized in Table 12.

The derivative of compound functions can be calculated using the following calculation rules:

Figure 15: Derivative of a function



Sum rule:

$$\frac{d}{dx}(f(x) \pm g(x)) = f'(x) \pm g'(x).$$

Multiplication with a number λ :

$$\frac{d}{dx}(\lambda f(x)) = \lambda f'(x).$$

Product rule:

$$\frac{d}{dx}(f(x) \cdot g(x)) = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

Quotient rule:

$$\frac{d}{dx} \left(\frac{f(x)}{g(x)} \right) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{[g(x)]^2}$$

Chain rule:

$$\frac{d}{dx}(F(g(x))) = F'(g(x)) \cdot g'(x).$$

Higher derivatives

If the first derivative is differentiated again, then this becomes the second derivative. Higher derivatives can be formed in a similar way.

2nd derivative:

$$\frac{d}{dx}f'(x) = f''(x) \text{ or } f''(x) = \frac{d^2}{dx^2}f(x)$$

3rd derivative:

$$\frac{d}{dx}f''(x) = f'''(x) \text{ or } f'''(x) = \frac{d^3}{dx^3}f(x).$$

n th derivative ($n > 3$):

$$\frac{d}{dx}f^{(n-1)}(x) = f^{(n)}(x) \text{ or } f^{(n)}(x) = \frac{d^n}{dx^n}f(x).$$

Extrema of functions

If the first and the second derivative of a function $f(x)$ are on hand, then these can be used to determine the maxima and minima (extrema) of $f(x)$. Both of the following conditions apply to an extremum at x_0 :

- $f'(x_0) = 0$,
- $f''(x_0) < 0$ for a maximum,
- $f''(x_0) > 0$ for a minimum.

Inflection point

The curvature properties of the function can be determined by using the second derivative. The point x_w at which a function switches from left-curved ($f''(x) > 0$, convex curvature) to right-curved ($f''(x) < 0$, concave curvature) and vice versa is of particular interest. These points are known as inflection points. The following is true here:

$$f'''(x) = 0.$$

Integration of a function

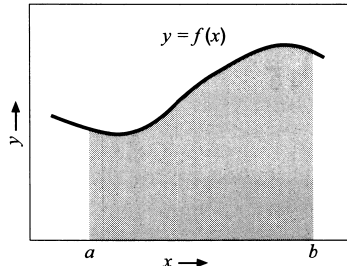
A (continuous) function $f(x)$ is depicted in Figure 16. Together with the X -axis in

Table 12: Elementary functions $f(x)$ and their derivatives $f'(x)$

$f(x)$	$f'(x)$
$c = \text{const}$	0
x^n	$n x^{n-1}$
$\sqrt{x} = x^{1/2}$	$\frac{1}{2} \frac{1}{\sqrt{x}} = \frac{1}{2} x^{-1/2}$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x)$	$\frac{1}{\cos^2(x)}$
$\cot(x)$	$-\frac{1}{\sin^2(x)}$
e^x	e^x
$a^x = e^{x \ln a}$	$(\ln a) a^x$
$\ln x$	$\frac{1}{x}$
$\log_a x = \frac{\ln x}{\ln a}$	$\frac{1}{\ln a} \frac{1}{x}$

Figure 16: Integral

Integration of the function $f(x)$ in the interval of a through b .



the interval $[a; b]$, this encloses the area A marked in gray. Its dimensions can be calculated through integration.

The unknown area A is represented as an integral from a to b via the function $f(x)$ with the variable x :

$$A = \int_a^b f(x) dx.$$

Antiderivative

In order to find a solution to this problem, the antiderivative $F(x)$ is required for $f(x)$. Its property is expressed in the fundamental theorem of differential and integral calculus.

$$F(x) = \int_a^x f(t) dt$$

is an antiderivative for $f(x)$, if the following applies:

$$\frac{d}{dx} F(x) = f(x).$$

If $F(x)$ is an antiderivative for $f(x)$, then the unknown area can be calculated as

$$A = \int_a^b f(x) dx = F(b) - F(a).$$

Important antiderivatives are compiled in Table 13. For each function $f(x)$, there is just one antiderivative $F(x)$. This is

Table 13: Elementary functions and their antiderivatives (examples)

$f(x)$	$\int f(x) dx$
0	$c = \text{const}$
$x^n, c \neq -1$	$\frac{1}{n+1} x^{n+1} + C$
$\frac{1}{x}$	$\ln x + C$
$\sqrt{x} = x^{1/2}$	$\frac{2}{3} x^{3/2} + C$
$\sin x$	$-\cos x + C$
$\cos x$	$\sin x + C$
$\frac{1}{\cos^2 x}$	$\tan x + C$
$\frac{1}{\sin^2 x}$	$-\cot x + C$
e^x	$e^x + C$
$a^x = e^{x \ln a}$	$\frac{a^x}{\ln a} + C$

uniquely defined with the exception of the integration constant C .

Calculation rules for integration

The following calculation rules apply to integration:

Sum rule:

$$\int_a^b f(x) + g(x) dx = \int_a^b f(x) dx + \int_a^b g(x) dx.$$

Multiplication with a number λ :

$$\int_a^b \lambda f(x) dx = \lambda \int_a^b f(x) dx.$$

Reversing the bounds of integration:

$$\int_a^b f(x) dx = - \int_b^a f(x) dx.$$

Lower parts of integration ranges in intervals ($a < c < b$):

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Upper and lower bounds of integration are identical:

$$\int_a^a f(x) dx = 0.$$

Partial integration

In addition to searching for the antiderivative, there are two helpful methods for solving for integrals for partial integration and integration through substitution. Partial integration is the inverse of the product rule for the derivative.

$$\begin{aligned} \int_a^b f'(x) \cdot g(x) dx &= \\ f(x) \cdot g(x) - \int_a^b f(x) \cdot g'(x) dx. \end{aligned}$$

The critical factor for this method is properly assigning both product terms $f'(x)$ and $g(x)$ within the integral.

Substitution

Substitution is the inverse of the chain rule. The art of substitution lies in finding the appropriate function that can be replaced (for more information, refer to [1], [2], [3]).

Linear differential equations

In many technical problems, a function is being sought that can describe the task but is not explicitly specified. Instead, information about its curvature properties (second derivative), its slope (first derivative) or a combination of the two is provided in the form of an equation.

An equation in which the function and its derivatives can be found is known as a differential equation. Equations known as common linear differential equations of the n th order with constant coefficients are covered below ([1], [2] and [3]):

$$y^{(n)}(x) + a_{n-1}y^{(n-1)}(x) + \dots + a_0y(x) = g(x).$$

In this equation, the n th derivative of $y(x)$ is the highest order derivative. Real numbers stand in front of individual derivatives $a_0, a_1, \dots, a_{n-1}, a_n$, where $a_n = 1$ is selected. If the function $g(x)$ is always zero, the differential equation is homogeneous; otherwise, it is non-homogeneous. Here, x is the variable of the unknown function $y(x)$.

The differential equation above is linear, since $y(x), y'(x), \dots, y^{(n)}(x)$ are only accompanied by the coefficients and do not occur as exponents, such as $y(x)^2$, or as a nonlinear function, like $\sin(y'(x))$.

Using the formulation

$y(x) = Ae^{\lambda x}$, $A \in \mathbb{R}$, $\lambda \in \mathbb{R}$, a homogeneous differential equation is simplified to a characteristic polynomial, because the exponential term is canceled out:

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_2\lambda^2 + a_1\lambda + a_0 = 0.$$

In doing so, solving the differential equation is reduced to searching for the zero points of the characteristic polynomial. If the n zero points $\lambda_1, \lambda_2, \dots, \lambda_n$ are determined, then the homogeneous solution $y_h(x)$ of the differential equation is yielded as

$$y_h(x) = A_1e^{\lambda_1 x} + A_2e^{\lambda_2 x} + \dots + A_ne^{\lambda_n x}.$$

The numbers $A_1, A_2, \dots, A_n$ are (any) real numbers and are labeled as integration constants. If initial or ancillary conditions are given for the function, then the values $A_1, A_2, \dots, A_n$ can be calculated from these conditions.

After finding the homogeneous solution, a special solution $y_{\text{inh}}(x)$ for the non-homogeneous differential equation needs to be found; this then provides the complete solution.

$$y(x) = y_h(x) + y_{\text{inh}}(x).$$

Linear differential equations of the second order with constant coefficients are the basis of all oscillation processes (see Vibrations and oscillations, Vibration characteristics) and also the basis for the design of safety-related sensors like acceleration and vibration sensors (see Acceleration sensors). Even regulating and control engineering technology is based on these differential equations (see Control engineering).

In addition to common linear differential equations with constant coefficients, there are also linear differential equations in which the constant coefficients are replaced by functions, such as

$$\sin(x) \cdot y'(x) + \cos(x) \cdot y(x) = 0.$$

Furthermore, the two-dimensional oscillation equation of the function $f(x, y)$ with the location coordinates x and y and time t ,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2}, \quad c \in \mathbb{R},$$

is an example of a partial differential equation. It is significantly more complicated to solve both of these types of differential equations than it is to solve the linear differential equations described above. For more information and a more detailed approach to these differential equations, the extensive literature from the bibliography should be consulted, including [1], [2] and [3], or the specialized literature regarding special types of differential equations.

Laplace transform

There are many control loops in a vehicle (see Control engineering); in the engine (e.g. knock control, λ control), in the A/C unit or in the undercarriage (e.g. yaw-rate control). These control loops are often represented by linear differential equations. One way to solve this differential equation is to use the exponential approach (see Differential equations).

Alternatively, the Laplace transform can also be utilized if initial values like $y(0)$, $y'(0)$ etc. are given in addition to the differential equation ([1], [2], [3]). The advantage to this approach is that the Laplace transformation converts the differential equation into an algebraic equation. This can usually be resolved easily according to the Laplace transform function $Y(s)$. Then, the Laplace inverse transform must be carried out in order to find the unknown function $y(x)$.

The Laplace transform $\mathcal{L}\{y(x)\}$ or image function $Y(s)$ of a function $y(x)$ can be found as follows:

$$Y(s) = \mathcal{L}\{y(x)\} = \int_0^{\infty} e^{-sx} y(x) dx, \quad s \in \mathbb{R}.$$

In order to solve for the integral, the absolute value of the function must be exponentially bound. s merely acts as a variable in the transformation [3].

Properties

Laplace transforms of the derivatives $y'(x)$, $y''(x)$ are very crucial when solving differential equations.

$$\mathcal{L}\{y'(x)\} = sY(s) - y(0)$$

$$\mathcal{L}\{y''(x)\} = s^2 Y(s) - sy(0) - y'(0)$$

$$\begin{aligned} \mathcal{L}\{c_1 y_1(x) + c_2 y_2(x)\} \\ = c_1 \mathcal{L}\{y_1(x)\} + c_2 \mathcal{L}\{y_2(x)\} \end{aligned} \quad (\text{linearity})$$

$$\mathcal{L}\{y(ax)\} = \frac{1}{a} Y\left(\frac{s}{a}\right), \quad a > 0$$

(Similarity theorem)

$$\mathcal{L}\{e^{-ax} y(x)\} = Y(s+a), \quad a > 0$$

(Complex shifting theorem)

As you can see, the derivatives are converted into terms with the function $Y(s)$.

A few selected examples of the Laplace transforms of functions have been compiled in Table 14.

Table 14: Functions $y(x)$ and the corresponding Laplace transforms $Y(s)$

$y(x)$	$Y(s) = \mathcal{L}\{y(x)\}$
$y(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$	$\frac{1}{s}$
$y(x) = \begin{cases} 0 & x < 0 \\ x^{n-1} & x \geq 0 \end{cases}$	$\frac{(n-1)!}{s^n} \quad n=1, 2, 3$
$y(x) = \begin{cases} 0 & x < 0 \\ \sin(ax) & x \geq 0 \end{cases}$	$\frac{a}{s^2 + a^2}$
$y(x) = \begin{cases} 0 & x < 0 \\ \cos(ax) & x \geq 0 \end{cases}$	$\frac{s}{s^2 + a^2}$
$y(x) = \begin{cases} 0 & x < 0 \\ e^{-ax} & x \geq 0 \end{cases}$	$\frac{1}{s+a}$

Fourier transform

During analysis, a time-dependent function is often found $f(t)$. Analysis is usually interested in the frequencies of the vibrations that underlie this function. For example, the undercarriage of a car should compensate for the bumps in the road so that the vehicle does not begin to rock back and forth. This requires a surface unevenness profile $h(x)$. If the car is driving down the road at a particular speed, it will begin to rock, which corresponds to the function $f(t)$. The characteristic frequencies must be determined from this function in order to be able to select and set the dimensions for the vehicle's shock absorbers, which will dampen the oscillations (see Undercarriage, Road excitation, Irregularity profiles).

The characteristic oscillation frequencies can be determined using the function $f(t)$, if this function $f(t)$ is subjected to a Fourier transform to form the function $F(\omega)$:

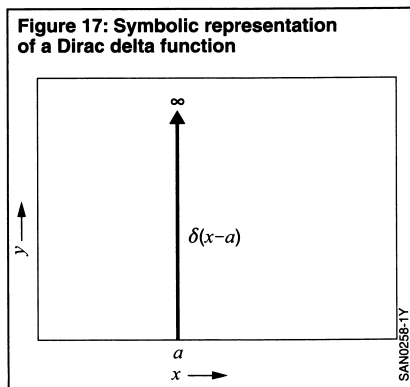
$$F(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt,$$

where $i^2 = -1$.

This converts the time-dependent function $f(t)$ into the frequency-dependent spectral function $F(\omega)$ ([1], [2]).

This results in an inverse transformation along the lines of:

$$f(t) = \mathcal{F}^{-1}\{F(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega.$$



In the literature, values of 1 (like here), $1/2\pi$ or $1/\sqrt{2\pi}$ are sometimes specified as the prefactor for the Fourier transform. This also changes the factor of the inverse transform. The critical aspect here is that the product of the transform's prefactors and the inverse transform yield $1/2\pi$.

Properties

Some properties of the Fourier transform are listed below.

$$\mathcal{F}\{c_1 f_1(t) + c_2 f_2(t)\} = c_1 \mathcal{F}\{f_1(t)\} + c_2 \mathcal{F}\{f_2(t)\} \quad (\text{Linearity})$$

$$\mathcal{F}\{f_1(at)\} = \frac{1}{|a|} F\left(\frac{\omega}{a}\right), \quad a \in \mathbb{R}, a \neq 0, \quad (\text{Similarity theorem})$$

$$\mathcal{F}\{f(t - t_0)\} = e^{-i\omega t_0} F(\omega), \quad t_0 \in \mathbb{R}, \quad (\text{Shift theorem in the time range})$$

$$\mathcal{F}\{e^{i\omega_0 t} f(t)\} = F(\omega - \omega_0), \quad \omega_0 \in \mathbb{R}, \quad (\text{Shift theorem in the frequency domain})$$

$$\mathcal{F}\left(\int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau\right) = \mathcal{F}\{f(t)\} \cdot \mathcal{F}\{g(t)\}.$$

Using the notation

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau \\ = \int_{-\infty}^{\infty} f(t - \tau) g(\tau) d\tau$$

the convolution theorem can be summarized as follows:

$$\mathcal{F}\{(f * g)\} = F(\omega) G(\omega).$$

The transform in the frequency domain can be explained by the Dirac delta function, which is defined as follows [4]:

$$\delta(t) = 0 \text{ for } t \in \mathbb{R}, t \neq 0, \text{ and}$$

$$\int_{-\infty}^{\infty} \delta(t - t_0) f(t) dt = f(t_0).$$

This can also be written as

$$\delta(t - t_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(t-t_0)\tau} d\tau.$$

The Dirac delta function is always zero. However, at the point where its argument is zero, it can be said to be infinitely large. This is illustrated in Figure 17. The Fourier transform of the complex exponential function or the sine and cosine function with an angular frequency of ω_0 returns the Dirac delta function for ω_0 :

$$\mathcal{F}\{e^{-i\omega_0 t}\} = 2\pi \delta(\omega - \omega_0) dx,$$

$$\mathcal{F}\{\delta(t - t_0)\} = \frac{1}{2\pi} e^{-i\omega t_0},$$

$$\mathcal{F}\{\sin(\omega_0 t)\} = i\pi [\delta(\omega + \omega_0) - \delta(\omega - \omega_0)],$$

$$\mathcal{F}\{\cos(\omega_0 t)\} = \pi [\delta(\omega + \omega_0) + \delta(\omega - \omega_0)].$$

For example, the Fourier transform is used to determine the frequencies of signals (e.g. the individual tones within the sound of a piano). In addition, the ratios of the corresponding amplitudes corresponding to the volume of sounds can be found. This allows sound waves to be broken down into individual sounds (frequencies) and their amplitude (volume) (see Undercarriage, Road excitation; Acoustics, Perceived sound levels).

Similarities and differences between Laplace and Fourier transforms

The Laplace transform of a function $y(x)$

$$\mathcal{L}\{y(x)\} = \int_0^{\infty} e^{-sx} y(x) dx, s \in \mathbb{R}, \mathbb{C}$$

and the Fourier transform of $y(x)$

$$\mathcal{F}\{y(x)\} = \int_{-\infty}^{\infty} e^{-i\omega x} y(x) dx, \text{ where } i^2 = -1$$

are similarly defined. This means that both transforms have properties that are either similar or the same. Both transforms are linear, their similarity theorems are almost identical and the complex shifting theorem of the Laplace transform is equivalent to the shift theorem of the Fourier transform.

However, both transforms can be distinguished by their bounds of integration, because the Laplace integral begins at $x = 0$ and the Fourier integral at $-\infty$. In other words, the function $y(x)$ can be defined by the entirety of $\mathbb{R}$ in the case of the Fourier integral, while $y(x)$ can only differ from zero for $\mathbb{R}^+$. Furthermore, the

exponential function of the Laplace integral usually contains an arbitrary complex exponent $-sx$, while this term $-i\omega x$ is normally completely imaginary in the case of a Fourier integral.

The inverse of the Fourier transform can usually be calculated in a manner similar to that of the Fourier transform. However, calculating the inverse Laplace transform by means of complex curvilinear integrals is often difficult [1]. Tables such as Table 14 are usually helpful when solving differential equations in order to determine the original function $y(x)$ from the Laplace transform.

The direct correlation between the Laplace and the Fourier transform is depicted as follows: For the function $y(x)$, the function $g(x)$ is defined as:

$$g(x) = \begin{cases} 0 & x < 0 \\ e^{-\lambda x} y(x) & x \geq 0 \end{cases} \quad \lambda \in \mathbb{C}$$

Subjecting $g(x)$ to a Fourier transform yields

$$\begin{aligned} \mathcal{F}\{g(x)\} &= \int_{-\infty}^{\infty} e^{-i\omega x} g(x) dx \\ &= \int_0^{\infty} e^{-i\omega x} e^{-\lambda x} y(x) dx. \end{aligned}$$

Solving for $s = \lambda + i\omega$ then yields

$$\begin{aligned} \mathcal{F}\{g(x)\} &= \int_{-\infty}^{\infty} e^{-i\omega x} g(x) dx = \int_{-\infty}^{\infty} e^{-sx} y(x) dx \\ &= \mathcal{L}\{y(x)\}. \end{aligned}$$

Matrix calculus

Vectors can indicate the intensity and direction of forces or of a velocity (see Vectors). In the case of a circular motion this can be easily represented using a vector, which specifies the initial direction of motion, together with a matrix, which describes the rotation.

Furthermore, matrices play the central role in solving systems of linear equations which occur in many technical questions. In order for example to calculate mechanical or thermal stresses in a component, complicated nonlinear differential equations are numerically solved. The standard method used for this purpose is the finite-element method (FEM), which ultimately (in very small time and space intervals) converts the differential equations into systems of linear equation and consequently into matrices. Reference is also made here by way of example to the use of matrices in the image processing of camera data in passenger cars (see Computer vision).

Firstly, matrices will be introduced in the following. Then arithmetical operations with matrices such as addition and multiplication will be defined. To conclude, three solution methods for systems of linear equation will be presented.

Structure of matrices

An $(m \times n)$ matrix A is made up of m rows and n columns. The component (elements, entries) a_{ik} in the i th row and k th column is a real or a complex number.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ik} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mk} & \dots & a_{mn} \end{bmatrix}$$

If the matrix has the same number of rows and columns ($m=n$), this is called a square matrix. The standouts among these are the upper and lower triangular matrices, the diagonal matrices, and the unit matrix. An upper or lower triangular matrix (A_o or A_u) has different entries on the diagonal and above or below zero;

below or above the diagonals there are only zeros.

$$A_o = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}, \quad A_u = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

A diagonal matrix D has different elements a_{kk} only on the diagonal. The unit matrix E is a diagonal matrix with ones as entries.

$$D = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}, \quad E = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Arithmetical rules for matrices

Two matrices A and B are identical when they match in all components:

$$A = B \Leftrightarrow a_{ik} = b_{ik} \text{ for all } i, k.$$

Two matrices can only then be added or subtracted when they match both in their number of rows and their number of columns. In this case, the addition (subtraction) is performed for each component:

$$C = A + B, \quad c_{ik} = a_{ik} + b_{ik}.$$

If a matrix is multiplied by a real or complex number λ , each individual components is multiplied by this scalar:

$$B = \lambda A, \quad b_{ik} = \lambda a_{ik}.$$

This results in the following arithmetical rules for matrices:

$$\begin{aligned} A + B &= B + A, \\ A + (B + C) &= (A + B) + C, \\ \lambda(\mu A) &= (\lambda\mu)A, \\ (\lambda + \mu)A &= \lambda A + \mu A, \\ \lambda(A + B) &= \lambda A + \lambda B. \end{aligned}$$

A column vector x with n entries can be multiplied from the right by an $(m \times n)$ -matrix as follows:

$$Ax = \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \dots a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}$$

If the rows of matrix A are considered as vectors, i.e. the k th row corresponds to the row vector $a_k = (a_{k1} \ a_{k2} \ a_{k3} \dots a_{kn})$, then in this multiplication in each case the scalar product is formed from the row vector a_k and the column vector x .

Thus, for the product of an $(m \times n)$ matrix with an n -row vector, an m -row vector is produced. This multiplication is linear, i.e. for matrix A , the two vectors x, y , and the two scalars λ, μ the following applies: $A(\lambda x + \mu y) = \lambda Ax + \mu Ay$.

Similarly, the multiplication of a row vector can be defined from the left with a matrix; this will not be discussed here.

Two matrices can be multiplied by each other when the number of columns s of the first matrix A matches the number of rows of the second matrix B . When the columns of matrix B are interpreted as vectors, the matrix multiplication is a natural extension of the multiplication of a matrix by a vector:

$$C = AB = \begin{bmatrix} a_{11} & a_{12} \dots a_{1s} \\ a_{21} & a_{22} \dots a_{2s} \\ \vdots & \vdots \\ a_{i1} & a_{i2} \dots a_{is} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \dots a_{ms} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \dots b_{1j} \dots b_{1n} \\ b_{21} & b_{22} \dots b_{2j} \dots b_{2n} \\ \vdots & \vdots \\ b_{s1} & b_{s2} \dots b_{sj} \dots b_{sn} \end{bmatrix} = \begin{bmatrix} c_{11} \dots c_{1j} \dots c_{1n} \\ c_{i1} \dots c_{ij} \dots c_{in} \\ \vdots \\ c_{m1} \dots c_{mj} \dots c_{mn} \end{bmatrix}$$

where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} \dots a_{is}b_{sj}.$$

Similarly to the above, the new element c_{ij} corresponds to the scalar product from the i th row vector $a_i = (a_{i1} \ a_{i2} \ a_{i3} \dots a_{is})$ of matrix A with the j th column vector $b_j = (b_{1j} \ b_{2j} \ b_{3j} \dots b_{sj})$ of matrix B . Thus, for the product of an $(m \times s)$ matrix with an $(s \times n)$ matrix an $(m \times n)$ matrix is produced.

The following arithmetical rules are obtained with regard to multiplication:

$$\begin{aligned} A(BC) &= (AB)C, \\ A(B+C) &= AB + AC, \\ (A+B)C &= AC + BC, \\ AE &= EA = A, \end{aligned}$$

It must be noted in this connection that when multiplying two matrices as a rule the order may not be mixed up (not commutative).

Determinants

Square matrices can be characterized by the so-called determinant. When the columns (or rows) of the determinant are interpreted as vectors, the value of the determinant indicates the volume of the parallelepiped spanned by these vectors. It is clear from this that the determinant is only then different from zero when the column or row vectors are linearly independent.

Notations for the determinants

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix}$$

For (2×2) and (3×3) matrices the determinant can be calculated using Sarrus' rule. Here all the elements which are on diagonals going from top left to bottom right are multiplied by each other. These diagonals are then added. The elements which are on diagonals going from top right to bottom left are also multiplied by each other and subtracted from the other values. To this end, it is appropriate particularly for (3×3) matrices to write the first two column vectors on the right next to the determinant to simplify the calculation.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} \\ + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \\ - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \end{matrix}$$

When the determinants have more than three rows (columns), the determinant can be calculated with the Laplace expansion.

For the unit matrix the following applies: $\det(E) = 1$.

For the product of two square matrices the following applies:
 $\det(AB) = \det(A) \cdot \det(B)$.

Inverse matrix

The square $(n \times n)$ matrix B is the inverse to the square $(n \times n)$ matrix A when the following applies:
 $AB = BA = E$.

In this special case the multiplication is commutative. Then $B = A^{-1}$ is written and the preceding equation becomes
 $AA^{-1} = A^{-1}A = E$.

The following is obtained for the determinant of an invertible matrix:
 $\det(A) \neq 0$ and $\det(A^{-1}) = \frac{1}{\det(A)}$

Systems of linear equations

In the case of a system of linear equations real (or complex) numbers a_{ij} and b_j ($j = 1 \dots m$) are given while the x_i ($i = 1 \dots n$) are unknown and thus are to be determined.

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \quad \quad \quad \dots \quad \quad \quad \dots \quad \quad \quad \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{matrix}$$

These equations can be summarized compactly with the aid of the $(m \times n)$ matrix A and the two vectors b, x

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

as a matrix-vector equation:
 $Ax = b$

Three methods which can be used to solve the system of linear equations are set out below.

1) In the case of a square and furthermore invertible matrix A the following is obtained:

$$x = A^{-1}b.$$

2) Alternatively, Cramer's rule can be used to solve a system of linear equations of a square and furthermore invertible matrix A . Here the component x_i of the solution vector x is obtained as:

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

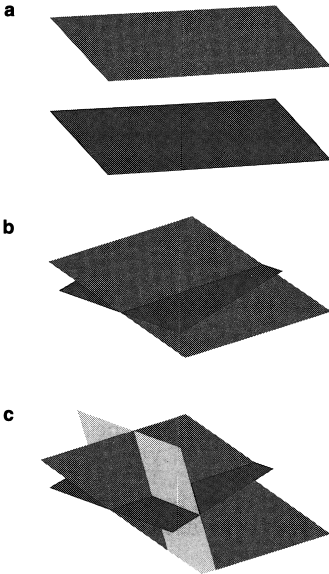
In the calculation of the auxiliary determinant A_i the i th column of the matrix is replaced by the vector b .

3) In the case of any $(m \times n)$ matrix the Gauss-Jordan algorithm can be used to determine the solution of the system of linear equations. Here matrix A is successively converted by row operations into an upper triangular matrix. The following are permitted as row operations:

- the multiplication of a row by a scalar (without zero), and
- the addition or subtraction of two rows.

Figure 18: Intersection of planes in space

- a) Parallel planes,
 b) Two planes intersect on a straight line,
 c) Three planes intersect at a point.



SAN0272Y

References

- [1] T. Arens et al.: Mathematik. 3rd Ed., Verlag Springer Spektrum, 2015.
 [2] I. N. Bronstein et al.: Taschenbuch der Mathematik. 10th Ed., Europa Lehrmittel, 2016.
 [3] L. Papula: Mathematik für Ingenieure und Naturwissenschaftler, Volumes 1–3. Verlag Springer Vieweg, 2014.
 [4] F. W. J. Oliver et al. (Editors): NIST Handbook of Mathematical Functions. 1st Ed., Cambridge University Press, 2010.

Geometric interpretation of a system of linear equations

Each row of a system of linear equations can be interpreted geometrically as a hyperplane. Thus, the approach when solving a system of linear equations is to look for the points where these hyperplanes intersect. This is clearly shown in Figure 1 by way of example at two or three planes in $\mathbb{R}^3$: When the planes are parallel to each other, they do not intersect. Two planes intersect on a straight line. The latter corresponds exactly to a solution of the system of linear equations.

Technical statistics

Several thousand parts are required to produce an automobile. Since (random or systematic) fluctuations occur in every production process, not only the size for example of each part but also the associated tolerance is specified.

Even the measurement process is as a rule subject to random fluctuations, which are referred to as measurement uncertainty (previously the misleading term measuring error was also used). This measurement uncertainty should however be well below the tolerance limit of the part to be measured so that it can normally be disregarded.

Measurements play a central role in production and hence in this connection in quality assurance. Since the required failure rate must be very low for individual parts, i.e. often in the ppm range (parts per million), it is for the most part not possible for process-technical and economic reasons to measure all the parts. Instead, samples are taken in order then to infer the quality of the components with statistical methods. These statistical methods are presented in the following.

The aim of a production process is to work accurately and precisely. Accurate means that the mean value of the target figure is close to the specified value. Precise denotes the production when the dispersion (scatter) around the mean value is low. It is possible to check using the statistical methods set out in the following whether these requirements are satisfied.

Table 1: Linear measurement and associated absolute frequency of the measured value

<i>l</i> [mm]	190	191	192	193	194	195
<i>h</i>	0	0	27	23	0	0
<i>l</i> [mm]	196	197	198	199	200	201
<i>h</i>	0	0	34	0	0	2
<i>l</i> [mm]	202	203	204	205	206	
<i>h</i>	6	1	0	0	0	

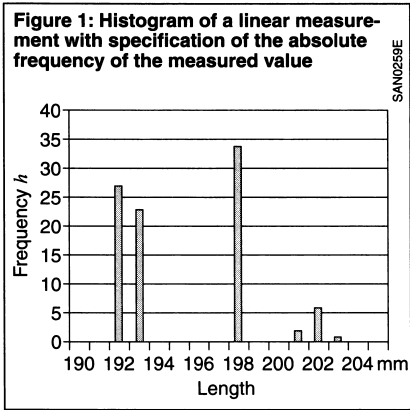
Descriptive statistics

For the purpose of obtaining an overview of extensive data material, these data are summarized and illustrated in descriptive statistics in tables and graphics. In this way, already systematic errors (or even typing errors) can often be detected and eliminated. However, it is not possible with descriptive statistics to specify error probabilities, which can only be made on the basis of statistical parameter models of inductive statistics.

Absolute frequencies of a measurement

The starting point is a data set with *n* values *x*₁, *x*₂, ..., *x*_{*n*} (e.g., linear measurements) on *n* different components. These measurements are summarized according to their characteristics, where for each length *l*_{*i*} that occurs the associated number (absolute frequency) *h*_{*i*} is compiled (*i* = 1, 2, ..., *k*; *k* ≤ *n*). This is shown by way of example for a linear measurement in Table 1 and in the associated histogram (Figure 1).

In this example the length *l* = 198 mm occurs with the most frequency. For *l* = 192 mm and *l* = 193 mm an accumulation occurs, as well as around *l* = 202 mm. Linear values *l* ≥ 204 mm and *l* ≤ 191 mm were not measured in this sample.



The sample size n corresponds to the number of measurements and results as the total of the absolute frequencies:

$$n = h_1 + h_2 + \dots + h_k = \sum_{i=1}^k h_i.$$

From this is obtained the relative frequency

$$f_i = \frac{h_i}{n} \text{ where } i = 1, 2, \dots, k.$$

When the relative frequency is added up, this produces the value 1:

$$\sum_{i=1}^k f_i = 1.$$

To characterize a data set and to be able to compare this with other data sets, characteristic values such as the location parameters and dispersion parameters are calculated from this.

Location parameters and location measures

The median x_{med} , also called the central value, is the value in the middle of a data series $\{x_1, x_2, \dots, x_n\}$ when the values are sorted by size. In other words, at least half of the data is less than or equal to x_{med} and at the same time at least half of the data is greater than or equal to x_{med} . In the case of an even number of data, x_{med} is the mean value of the two data in the middle of the series.

Often also of interest is the location of the lower (or upper) portion of values. The associated threshold value $x_{pQ} \in \{x_1, x_2, \dots, x_n\}$ is also referred to as p quantile or as percentile. Hence, therefore, at least $p \cdot 100\%$ of the data is less than or equal to x_{pQ} . At the same time, at least $(1-p) \cdot 100\%$ of the data is greater than or equal to x_{pQ} . The 25% quantile is also called the first or lower quartile; the 75% quantile the upper or third quartile.

The arithmetic mean (average value) is defined as

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

and the geometric mean of non-negative numbers x_i as

$$\bar{x}_{\text{geo}} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}.$$

The following always applies: $\bar{x}_{\text{geo}} \leq \bar{x}$.

Likewise, the arithmetic and geometric means can also be calculated via the characteristics l_i and their frequency h_i .

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k l_i h_i,$$

$$\bar{x}_{\text{geo}} = \sqrt[n]{l_1^{h_1} l_2^{h_2} \dots l_k^{h_k}}.$$

The symbol μ is often used as well as the symbol $\bar{x}$ for the mean value.

Dispersion parameters

In a data set it is also important to know whether the individual values are close to or far from the mean value. This is recorded by the following dispersion parameters. The variance $\text{Var}(x) = s^2$ of the data set x is calculated as

$$\begin{aligned} s^2 = \text{Var}(x) &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= \frac{1}{n-1} \sum_{i=1}^k h_i (l_i - \bar{x})^2. \end{aligned}$$

The root from this is called standard deviation s .

$$s = \sqrt{s^2} = \sqrt{\text{Var}(x)}$$

σ is often used as well as s for the standard deviation.

Example

For the above example (Figure 1 and Table 1) the mean value is obtained at $\bar{x} = 195.40$ mm. The two frequencies $h_{192 \text{ mm}}$ and $h_{193 \text{ mm}}$ added up produce a number 50. This is above $n/2 = 46.5$. Thus, the median follows with $x_{\text{med}} = 193$ mm. The variance has the value $s^2 = 11.48 \text{ mm}^2$ and the standard deviation $s = 3.38$ mm.

Probability calculus

The quality of a production process is to be inferred from a sample (e.g., linear measurement of a certain quantity of components). This is however only possible if realistic assumptions about the distribution of this measured value are made. A number of probability distributions are presented in the following for this purpose.

The Binomial distribution, the Hypergeometric distribution and the Poisson distribution are discrete distributions. Here, a sample is taken from a quantity and then the corresponding probabilities are calculated. From a mathematical standpoint both the sample size and the size of the quantity from which the sample is taken are natural numbers.

On the other hand, the Weibull distribution and the Gaussian normal distribution apply to continuous distributions, i.e. to real variables t or x .

Discrete distributions

Binomial distribution

A sample delivers only two values (Bernoulli experiment), one value with the probability p and the second value with the probability $q=1-p$ (since $p+q=1=100\%$). An example in this connection is drawing a ball from a bowl containing balls of two different colors. The ball is returned to the bowl after each draw. Hence the probability for each removal of a sample is the same.

The associated probability function $f(x)$ indicates that this event in the case of n -fold execution occurs exactly x times:

$$f(x) = \binom{n}{x} p^x q^{n-x} = \binom{n}{x} p^x (1-p)^{n-x},$$

where $x = 0, 1, 2, \dots, n$; $\binom{n}{x} = \frac{n!}{x! (n-x)!}$.

The binomial distribution has the mean value $\mu = np$, the variance $\sigma^2 = npq$ = $np(1-p)$, and the standard deviation $\sigma = \sqrt{npq} = \sqrt{np(1-p)}$.

Hypergeometric distribution

This distribution often plays a role in quality and final inspections. Similarly to the previous bowl model, there are at the start m white and $n-m$ red balls in a bowl. In contrast to the binomial distribution, this distribution corresponds to drawing a ball without putting it back, was which is customary for a sample in production. Hence, with each sample removal, the probability of drawing a white ball changes.

If now k samples are removed, this produces the probability function $f(x)$ at:

$$f(x) = \frac{\binom{m}{x} \binom{n-m}{k-x}}{\binom{n}{k}}, \quad x=0, 1, 2, \dots, k.$$

The hypergeometric distribution has the mean value

$$\mu = k \frac{m}{n}, \quad \text{the variance}$$

$$\sigma^2 = \frac{k m (n-m) (n-k)}{n^2 (n-1)}$$

and the standard deviation

$$\sigma = \sqrt{\frac{k m (n-m) (n-k)}{n^2 (n-1)}}.$$

For $k < 0.05n$ the easier-to-calculate binomial distribution can be used as a very good approximation for the hypergeometric distribution.

Poisson distribution

If a binomial distribution exists with a very low probability of occurrence p , this can also be described by a Poisson distribution with the mean value μ and the variance $\sigma^2 = \mu$:

$$f(x, \mu) = \frac{\mu^x}{x!} e^{-\mu}, \quad x = 0, 1, 2, \dots$$

The Poisson distribution has the mean value μ , the variance $\sigma^2 = \mu$, and the standard deviation $\sigma = \sqrt{\mu}$.

A binomial distribution can be well approximated by a Poisson distribution if the following two conditions are satisfied:

$$np < 10, \\ np > 1500p.$$

Continuous distributions

The Weibull distribution and the Gaussian normal distribution are classed as the continuous distributions. Hence they are functions of real variables x . The following functions are used to describe these: compression functions $f(x)$ and the associated distribution functions $F(x)$, which constitute their antiderivatives:

$$\frac{d}{dx}F(x) = f(x).$$

The integral over the compression function of x_1 through x_2

$$\int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1)$$

specifies the probability that the random variable x is in the interval of x_1 through x_2 . Accordingly, the distribution function

$$F(x_2) = \int_{-\infty}^{x_2} f(x) dx$$

is the probability that an event with $x \leq x_2$ occurs.

Often the expectation value $E(x)$ is referred to instead of the mean value μ

$$\mu = E(x) = \int_{-\infty}^{\infty} x f(x) dx.$$

The variance is defined as

$$\sigma^2 = \text{Var}(x) = E((x - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

Weibull distribution

The fatigue of a material and the failure of a machine can often be described via a Weibull distribution over the time t (instead of the variable x) (service life T of the material or the machine, empirical or specific parameters α, β).

Accordingly, the probability that the service life $\leq t$ is at

$$F(t) = 1 - e^{-at^\beta}$$

and hence the survival probability (reliability)

$$R(t) = 1 - F(t) = e^{-at^\beta}.$$

The associated compression function $f(t)$ is the first derivation of $F(t)$:

$$f(t) = F'(t) = \alpha \beta t^{\beta-1} e^{-at^\beta} \\ \text{for } t > 0, \text{ otherwise } f(t) = 0.$$

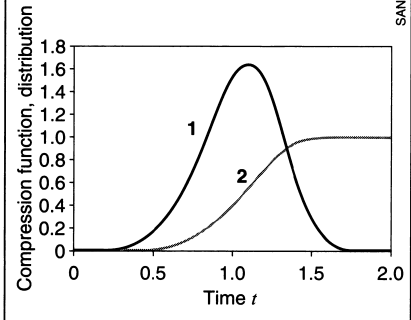
From this the failure rate is obtained at

$$\lambda = \frac{f(t)}{R(t)} = \alpha \beta t^{\beta-1}.$$

Figure 2 shows a compression function and the associated distribution for $\alpha=0.5$ and $\beta=5.0$. The greatest failure probability is at $t=1.1$. With increasing time the failure probability does in fact decrease again, but the function $F(t)$, i.e. the total failure rate until time t , naturally approaches $1 = 100\%$.

Figure 2: Weibull distribution

For $\alpha=0.5$ and $\beta=5.0$.
1 Compression function,
2 Distribution.



Gaussian normal distribution

Very many random processes in nature and also in technology can be described by a Gaussian normal distribution. The associated compression function contains the expectation value μ and the standard deviation σ :

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty.$$

The associated Gaussian normal distribution function

$$\begin{aligned} F(x, \mu, \sigma) &= \int_{-\infty}^x f(z, \mu, \sigma) dz \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(z-\mu)^2}{2\sigma^2}} dz, -\infty < x < \infty \end{aligned}$$

is the antiderivative for $f(x, \mu, \sigma)$. As the name already suggests, this is normalized, i.e.

$$\lim_{x \rightarrow \infty} F(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(z-\mu)^2}{2\sigma^2}} dz = 1.$$

Figure 3 shows by way of example a Gaussian compression function and the associated Gaussian normal distribution. The compression function has its maximum at the expectation value; its breadth is determined by the standard deviation. Furthermore, it is axisymmetric to $x = \mu$. This “Gaussian bell-shaped curve” drops markedly for large x values and for small x values and approaches asymptotically the x axis.

The Gaussian normal distribution increases from 0 (for $x \rightarrow -\infty$) to the value 1 (for $x \rightarrow \infty$). At $x = \mu$ it has the value $F(x = \mu) = 0.5$.

The Gaussian normal distribution has the mean value μ , the variance σ^2 , and the standard deviation σ .

With the substitution $y = \frac{x - \mu}{\sigma}$ the following standard normal distribution is obtained

$$\varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}, -\infty < z < \infty$$

Accordingly, the following normal distribution is obtained from this

$$\Phi(y) = \int_{-\infty}^y \varphi(z) dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{1}{2}z^2} dz, -\infty < y < \infty.$$

Many production processes are characterized by way of their σ level (Table 2). In the case of the 3σ level, for instance, the probability that the produced parts are defect-free is 93.3 %.

The 6σ level (Six Sigma Level) plays a central role in quality management. Departing from the above definition of the compression function, long-term mean shifts are included here. Thus, the 6σ level corresponds to 3.4 DPMO (Defects Per Million Opportunities), i.e. $3.4 \cdot 10^{-6}$ of the parts are outside the tolerance band.

For a normally distributed random variable X with the mean value μ and the standard deviation σ the following probabilities are obtained:

Figure 3: Gaussian normal distribution

- 1 Compression function with the expectation value $\mu = 0.5$ and the standard deviation $\sigma^2 = 1$.
- 2 Associated normal distribution for the compression function.

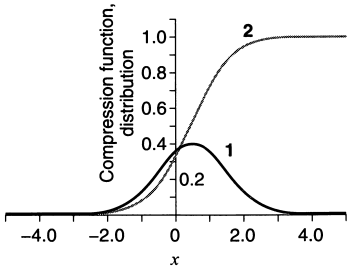


Table 2: σ level

The middle column specifies the probability that the values are inside the range $\mu - k\sigma < x < \mu + k\sigma$.

σ level	inside [%]	outside [%]
σ	32	68
2σ	69	31
3σ	93.3	6.7
4σ	99.38	0.62
5σ	99.977	0.023
6σ	99.99966	0.00034

$$P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \Phi(y).$$

$$P(X > x) = 1 - P(X \leq x) = 1 - \Phi(y).$$

$$P(a \leq X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

$$P(|X - \mu| \leq k\sigma) = P(\mu - k\sigma \leq X \leq \mu + k\sigma) = 2\Phi(k) - 1.$$

Parameter estimation in the Gaussian normal distribution

Quantitative statements based on samples are to be made with the aid of statistical models (distribution models). The aim would be to determine characteristic values such as, for example, the mean value μ and the standard deviation σ . But this is not exactly possible since the sample is limited. However, it is possible to calculate from the sample estimates for the mean value and the standard deviation and for this to determine confidence intervals with which the sought-after characteristic value is in this interval with a specified probability.

To this end a random sample is taken from the sample size n , delivering the values $x_1, x_2, \dots, x_n$. These elements are also called the realization of the sample. Calculated from these are the estimate of the mean value

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

and the estimate of the variance

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Parameter estimations are carried out in the following, where it is assumed that there is an underlying Gaussian normal distribution. This distribution underlies very many production processes. For the parameter estimations for other distributions, reference is made to the bibliography ([3], [4], [5]).

Confidence interval for unknown mean value and known variance

In the case of a specified confidence level γ (usually 95 % or 99 %) the confidence interval of the breadth Δx in which the unknown mean value μ lies with a probability of γ is to be determined, i.e. $P(|\bar{x} - \mu| \leq \Delta x) = \gamma$.

This is equivalent to the unknown mean value μ lying in the interval

$$\bar{x} - \Delta x \leq \mu \leq \bar{x} + \Delta x$$

with a probability of $\gamma \cdot 100$ %. The estimate of the mean value is calculated from the sample.

With the standardized random variable

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \text{ and } \Delta z = \frac{\Delta x}{\frac{\sigma}{\sqrt{n}}}$$

the above probability can be depicted as follows:

$$P(|z| \leq \Delta z) = \gamma.$$

From

$$P(|z| \leq \Delta z) = 2\Phi(\Delta z) - 1 \\ \Rightarrow \Phi(\Delta z) = 0.5(\gamma + 1)$$

the right side can be calculated and then Δz for example viewed in a table of the normally distributed Gaussian function, and finally from this Δx calculated.

Confidence interval for unknown mean value and unknown variance

From the sample first the estimates of the mean value $\bar{x}$ and the standard deviation s are calculated. Similarly to before, in the case of a specified confidence level γ (usually 95 % or 99 %) the confidence interval of the breadth Δx in which the unknown mean value μ lies with a probability of γ is to be determined, i.e.

$$P(|\bar{x} - \mu| \leq \Delta x) = \gamma.$$

The procedure for determining Δx is performed exactly as before, where for σ now s is used here.

Sums and products of random variables

If a process is determined by a number of random variables $x_1, x_2, \dots, x_n$, this produces for the sum total $z_S = x_1 + x_2 + \dots + x_n$ and for the product $z_P = x_1 \cdot x_2 \cdot \dots \cdot x_n$ the expectation values E and the variance:

$$E(z_S) = E(x_1) + E(x_2) + \dots + E(x_n)$$

$$\text{or } \mu_{z_S} = \mu_1 + \mu_2 + \dots + \mu_n$$

$$E(z_P) = E(x_1) \cdot E(x_2) \cdot \dots \cdot E(x_n)$$

$$\text{or } \mu_{z_P} = \mu_1 \cdot \mu_2 \cdot \dots \cdot \mu_n$$

$$\text{Var}(z_S) = \text{Var}(x_1) + \text{Var}(x_2) + \dots + \text{Var}(x_n)$$

$$\text{or } \sigma_{z_S}^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$$

Linear regression

In the case of a sample a linear connection between the parameter x and the measured variable y is presumed:

$$y = mx + t$$

The sample delivers for the set parameters x_i the measured values y_i ($i = 1, 2, \dots, n$). From these values the gradient of the straight line m and the axis intercept t are to be determined so that the deviation between the straight line and the measured values (x_i, y_i) is as small as possible (minimum Gaussian deviation). These are obtained then as

$$t = \frac{\left(\sum_{i=1}^n x_i \right) \cdot \left(\sum_{i=1}^n x_i y_i \right) - \left(\sum_{i=1}^n x_i^2 \right) \cdot \left(\sum_{i=1}^n y_i \right)}{\left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2}$$

and as

$$m = \frac{\left(\sum_{i=1}^n x_i \right) \cdot \left(\sum_{i=1}^n y_i \right) - n \sum_{i=1}^n x_i y_i}{\left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2}$$

In the case of a linear connection $y = mx$ without axis intercept, m is simplified at

$$m = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

Characteristic value in quality assurance and in quality management

Statistical Process Control (SPC) is used in production processes and in quality management to assure the quality of the manufactured goods. A Gaussian normal distribution is assumed as a rule here.

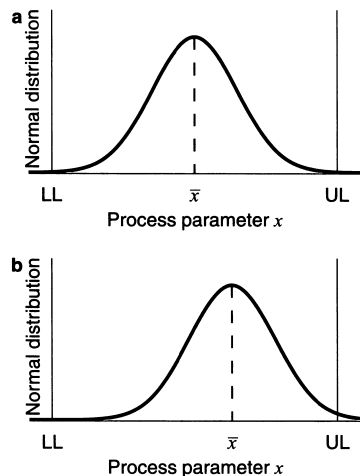
In a production process the upper and lower limits (UL and LL) are stipulated for example by customer requirements. The tolerance range T is obtained from this as

$$T = \text{UL} - \text{LL}$$

The mean value $\bar{x}$ and the standard deviation s are calculated from a sample. Furthermore, the critical length Z_{crit} is determined, i.e. the amount of the difference between the mean value $\bar{x}$ and UL or between the mean value $\bar{x}$ and LL is formed, and the smaller of these is taken as Z_{crit} .

Figure 4: Gaussian distributions of two samples

- a) Mean value lies exactly in the middle of the lower and upper limits,
 - b) Mean value is closer to the upper limit.
- $\bar{x}$ Mean value,
LL Lower limit,
UL Upper limit.



$$Z_{\text{crit}} = \min(|UL - \bar{x}|, |LL - \bar{x}|).$$

It can be seen in Figure 4 that the mean value is closer to the UL than to the LL. Hence here

$$Z_{\text{crit}} = UL - \bar{x}.$$

From this the process capability c_p is obtained as

$$c_p = \frac{T}{6s}$$

and the process capability characteristic value c_{pk} as

$$c_{pk} = \frac{Z_{\text{crit}}}{3s}.$$

In standard production $c_p \geq 1.33$ and $c_{pk} \geq 1.33$ are classed as guide values for a stable process which satisfies the quality requirements.

References

- [1] T. Arens et al.: Mathematik. 3rd Ed., Verlag Springer Spektrum, 2015.
- [2] I.N. Bronstein et al.: Taschenbuch der Mathematik. 10th Ed., Europa Lehrmittel Verlag, 2016.
- [3] L. Papula: Mathematik für Ingenieure und Naturwissenschaftler, Volumes 1 to 3. 14th Ed., Verlag Springer Vieweg, 2014
- [4] A. Steland: Basiswissen Statistik. 3rd Ed., Verlag Springer Spektrum, 2013.
- [5] G. Bamberg, F. Baur, M. Krapp: Statistik. 17th Ed., Oldenbourg Verlag, 2012.

Finite-element method

Applications

Virtually all technical procedures can be simulated on a computer with the finite-element method FEM (the term having been introduced by Ray W. Clough at the start of the 1960s [1]). However, this involves breaking down any body (gaseous, liquid, or solid) into elements that are simple in shape (line, triangle, square, tetrahedron, pentahedron, or hexahedron), that are as small as possible, and that are permanently connected to each other at their corner points ("nodes"). Small elements are important because the behavior of elements formulated by approximation using linear equations is only applicable to infinitesimal elements. However, the computing time calls for finite elements. The approximation to reality is better the smaller the elements are.

The application of FEM in practice – also known as FEA (finite-element analysis) – began in the early 1960s in the aviation and aerospace industries and followed soon after in automobile manufacturing. Today the method is used in all fields of technology, including weather forecasting, medical science, and for many sectors of automobile manufacturing ranging from engine and chassis components through to body calculations and crash behavior.

There are two different types of application. Firstly, virtually fully automatic "Black Box" FEM contained in all CAD programs (computer-aided design) for rough calculations by the design engineer (e.g. in designing a bumper) and, secondly, the use reserved for specialists of special FEM programs (e.g. in body calculations, in axle development or in driving dynamics).

FEM program system

The software of an FEM system consists of a preprocessor, a postprocessor, and the actual FEM program. Network creation, i.e. breakdown into elements, is mainly performed in the preprocessor on the basis of a CAD geometry which is read directly or via neutral interfaces such as IGES (Initial Graphics Exchange Specification),

VDA-FS (Verband der Automobilindustrie – Flächenschnittstelle) or STEP (Standard for the Exchange of Product Model Data). The FEM program calculates the computing model formulated in this way. The result found is then shown in graphic form in the postprocessor (e.g. stress distribution by means of isocolors, deformations as motion animation).

Basic knowledge for application

FEM is, like all numerical methods, an approximation process. In mechanics, the main area of application, the limitations caused by this are described in the following.

Small motions in one solution step

Bodies move on paths which are normally higher-order curves. With the basic principle of linearization of all processes, this motion is limited to a straight path which can then be described by linear equations. When transferred to the element corners (nodes), they also move on a straight line. Thus, the nodes are only able to realize very small motions correctly (node twists less than 3.5°). The actual motion along any path or nonlinear material behavior is thus solved linearly with many small steps.

Calculation accuracy

The linear equation system is formulated and solved with the limited computing accuracy of a computer. Usually, 8 bytes (= 64 bits) are used with a computing accuracy of 13 significant digits for the number stored, i.e. only the first 13 digits of a number can be represented exactly. The further digits in this number are random numbers. As a result, this rules out the possibility of any stiffness differences of the individual components in a model. Therefore, in the deformation calculation of a body, it is necessary, as for example in the measurement of body deformation, to replace the axle springs with rigid supports.

Interpretation of results

The great danger lies in the fact that a formal computation model formulated correctly by a beginner will indeed deliver beautifully colorful images, but the results shown can center around factors far from reality.

The problems resulting from the above-mentioned limitations must therefore be identified and signaled by the computing program, so that the less experienced user is also able to obtain correct results easily.

Areas of FEM application

In technology terms, physics is generally divided into five areas – mechanics with statics and kinematics (e.g. body, axle), dynamics with acoustics (e.g. vehicle noise), thermodynamics (e.g. temperature distribution in the engine), electricity with magnetism (e.g. ignition coil, sensor technology), and optics (e.g. headlamp). When it comes to FEM, a distinction is always made between

- linear and nonlinear static and dynamic problems with the deformations as an unknown for stress calculation and dynamic analysis,
- stationary (timeless) and nonstationary (time-dependent) potential problems (e.g. temperature, sound pressure, electrical or magnetic potential) with the potentials as an unknown,
- and the linking of these different fields, e.g. to calculate a temperature field and the resulting deformations, stresses and forces in linear statics when the engine is started.

Elements of FEM

The properties of the elements define the most important performance data of an FEM program. The element quality is determined by the degree of the mathematical formulation function selected. Here, for example, a distinction is made between elements with a linear or quadratic formulation, recognizable from the midside node in the middle of the edge. The quality of a computing model is therefore dependent not just on the fineness of the used mesh, but also quite considerably on the formulation function.

A distinction is made between three different types of elements: line elements, shell elements, and volume elements.

Line elements

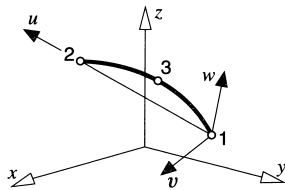
Line elements (Figure 1) are either straight or curved with an midside node. The cross-sections are described by specifying the numerical values for the cross-sectional area A , the reduced shear cross-sections $A_{\text{red-v-w}}$ (shear areas), the principal moments of inertia (I_v , I_w), the torsional moment of inertia (I_t) with the torsional section modulus (W_t), the sector moment of inertia for warping-force torsion, an angle α which describes the position of principal axes of inertia v and w in relation to the road plane, and the maximum four stress points (S_v , S_w) for bending-stress calculation.

Shell elements

Shell elements (Figure 2) are either triangular or quadrangular in shape – ideally an equilateral triangle or a square, usually of constant thickness. If the midside nodes are omitted, the edges are straight.

Figure 1: Line elements

With constant or linearly variable cross-section.



Volume (solid) elements

Volume elements (Figure 3) as tetrahedrons, pentahedrons and hexahedrons without midside nodes have straight edges – ideally an equilateral tetrahedron or a cube. With a sufficient number of elements in relation to the thickness (greater than or equal to three), elements with midside nodes are not necessary today. However, this does not apply to tetrahedrons, which are almost always used in the case of a complicated geometry, such as a cylinder head for example.

Modeling and evaluation of results

The most important function in using an FEM program is the usually time-consuming task of creating the input data as a computing model with the preprocessor. The user should try to achieve this target with as few elements and nodes as possible (however, the computing model of a body today has approximately three to

four million nodes). To do so, the user requires a certain level of experience, but must have an exact knowledge of the qualities of the elements used (see FEM examples). These can differ slightly in every FEM program.

The first step in modeling involves choosing the element type (line, shell or volume element) and determining the fineness of the mesh, e.g. by means of the specified middle element-edge length. The next step involves defining the properties (material data), e.g. the element thickness for unit elements, the cross-section values for line elements, and the units used (e.g. length in mm and force in N). A further step involves determining support conditions and load. The crucial factor here is considering the points where a model is fixed, and where it is loaded. In relation to loading, it is also useful to conduct a breakdown of the total load into load cases, e.g. into the mass load from gravity weight and different traffic loads.

All FEM results are available in list form or postprocessor data format, and can therefore be shown in graphic form (see FEM examples). To this end, the postprocessor offers all the conceivable forms of display.

Figure 2: Shell elements

a) Triangle, b) Quadrangular.

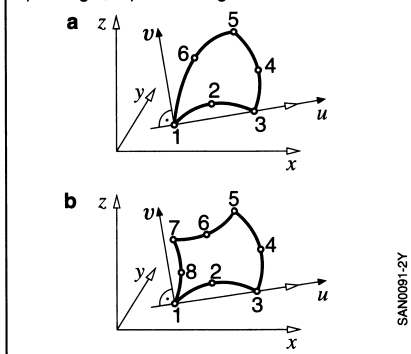
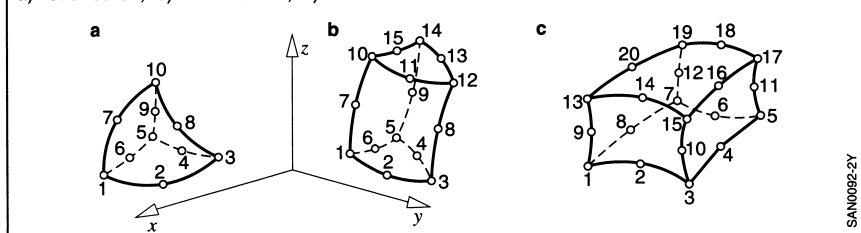


Figure 3: Volume elements

a) Tetrahedron, b) Pentahedron, c) Hexahedron.



FEM examples

For the examples, modeling is performed on the basis of CAD geometry using the FEM program TP2000. The model inputs, together with color representation of the results, can also be found on the Internet page specified in [2].

In reality, all bodies are three-dimensional. In simulation, a simplified solution is often chosen in order to save time and expense. It is much easier, for example, to realize the automatic meshing of a flat surface in shell elements than it is to realize a body in its volume elements. The frequently used tetrahedron mesh, which today every preprocessor creates for any

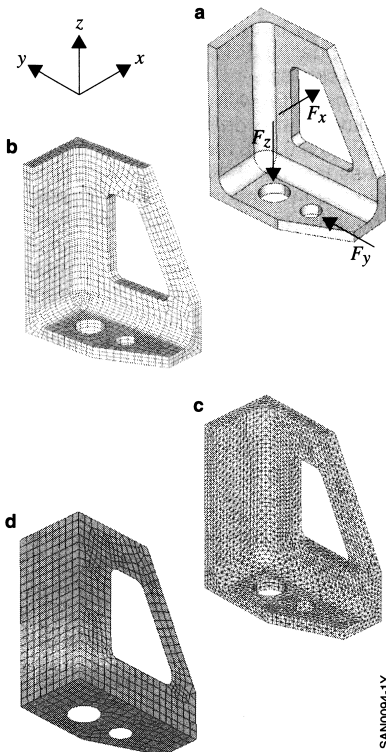
volume geometry, does not always live up to expectations.

In automobile manufacturing, the structural components are either thick-walled and solid (e.g. engine, transmission, axles, wheels) and modeled with volume elements or they are made from thin-walled sheet metal (e.g. automobile body, truck cab) and are modeled with shell and line elements.

The first example, as a volume-element model, belongs to the first group of thick-walled, solid components. It is compared with the commonly used shell model of the second group. This also includes the line element commonly used in automobile manufacturing with the second example.

Figure 4: Cast-steel engine mount

- a) CAD model with load F_x , F_y , F_z ,
- b) Model A1 (hexahedron),
- c) Model B1 (tetrahedron),
- d) Model C (shell).



Example 1: Cast-steel engine mount as shell- and volume-element model (linear statics)

The aim is to compare shell and volume elements in linear statics on a relatively thick-walled (3.75 mm) cast-steel engine mount ($50 \times 25 \times 57 \text{ mm}^3$). In the volume elements, two models A and B (each with and without midside nodes) with significantly different result qualities are compared; added to this is shell model C (Figure 4). Starting out from the CAD volume geometry, the preprocessor automatically creates volume-element networks A and B, and also shell-element model C by means of the surface model.

The material properties in the units mm, N and kg are:

Modulus of elasticity: 210,000 N/mm²

Poisson's ratio: 0.3

Density: 0.00000785 kg/mm³.

In the case of models A and B, the element properties are defined with the element type "Solid" (volume) with three node degrees of freedom v_x , v_y and v_z . In the case of model C, the properties are defined with the element type "Plate" (shell) with six node degrees of freedom v_x , v_y , v_z , d_x , d_y , d_z and constant thickness $d = 3.75 \text{ mm}$.

In model A, the meshing with volume elements (solid mesh) shows the preferred breakdown into hexahedrons (not possible fully automated for all geometries because here, for example, the geometry

first had to be broken down into basic bod-ies) and in model B, the automatic tetra-hedron meshing. The number of elements in relation to thickness which is crucial for accuracy is specified (at least three ele-ments here).

Support conditions

On the rectangular xz plane, $v_y = 0$ applies to all nodes. All the edge nodes on the right long leg are pinned by $v_x = 0$, and those of the lower short leg are pinned by $v_z = 0$.

Loading

$\Sigma F_x = 900\text{ N}$, $\Sigma F_y = 2006\text{ N}$, $\Sigma F_z = -550\text{ N}$. All the loads are defined as surface load ("on surface", for shell "on curve") at the recess at $F_x = 600\text{ N}$, at the small hole at $F_y = 2006\text{ N}$, and at the large hole at $F_z = -550\text{ N}$.

Note: "Isolated" individual loads are only permitted with line elements.

Result

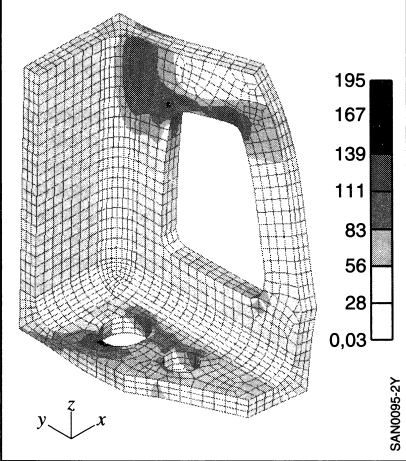
The results are shown in Table 1. Figure 5 shows the result for model A with stresses as shades of gray (critical areas are shown in dark) or as isocolors in the original.

Volume elements are very sensitive to incorrectly approximated load distribu-tions. Volume elements should therefore only ever be subjected to surface loads.

Considering Table 1 in detail: Experi-ence shows that model A2 with mid-side nodes delivers the correct result ($v = 0.064\text{ mm}$, $\sigma = 195\text{ MPa}$). The deviation between averaged and maximum node stress (from all the elements at the same node) should be as small as pos-

Figure 5: Deformation and stress distribution in model A2 as the correct model

$v = 0.064\text{ mm}$, $\sigma = 195\text{ MPa}$.



sible (less than 15 %). This is achieved with model A1 at 11 % with three elements per thickness.

In model B1 with three tetrahedrons per thickness, the deformation is 25 % too small ($v = 0.048\text{ mm}$), and the maximum stress is also 25 %. With only two tetra-hedrons in relation to the thickness, the model would even be 58 % too rigid and can barely be used with correspondingly lower stresses. However, with midside nodes (B2), it delivers virtually identical deformations and stresses to A2 with a very large number of nodes and long com-puting times. Caution is therefore advised with rough tetrahedron mesh without mid-side nodes.

Table 1: Results for engine-mount model

Values in parentheses apply to elements with midside nodes.

Model	Element type	No. of nodes	Weight in kg	Max. deformation in mm	Mean stress in MPa	Max. stress in MPa	Stress error max. in %	Computing time in sec
A1 (A2)	Solid hexahedron	4,000 (15,000)	0.119	0.061 (0.064)	145 (185)	164 (195)	15 (0)	10 (110)
B1 (B2)	Solid tetrahedron	10,000 (70,000)	0.119	0.048 (0.064)	115 (173)	146 (209)	30 (6)	60 (2,400)
C	Rectangular shell	766	0.114	0.081	197	223	14	2

Model C with shell elements without shear deformation (only intended for thin-walled structures) is clearly too soft ($\nu = 0.081$ mm, 21 % greater, as thick-walled with shear deformation even 30 % greater). Shell elements, whether thin- or thick-walled, deliver only partially usable deformations with relatively large thickness, particularly in the case of very compact bodies, as is the case here. However, at stresses of +14 %, it is for the most part on the safe side.

Example 2: Tubular body frame

The aim is to carry out an FEM analysis on a body in the form of a spaceframe without sheet paneling, including optimized weight and stiffness for the example of a mini-pickup (not a real vehicle). The material properties are:

Modulus of elasticity: 200,000 N/mm²

Poisson's ratio: 0.3

Density: 0.0000785 kg/mm³.

For reasons of simplicity, only two shapes are used as profile sections (Figure 6). A box-type profile section ($90 \times 120 \times 1.5$ mm³) and a tube profile section (70×2 mm²), which are defined by their shape in the preprocessor in the properties directly with the element type "Bar" and its constant cross-section (different starting and end cross-sections would be "Beam"). This results in the required cross-section values as follows (in mm² or mm⁴ and mm³).

Box/tube (Figure 6)

– Cross-sectional area:

$$A = 621/427.$$

– Reduced shear cross-sections (shear area):

$$A_{\text{redl}} = 325/227,$$

$$A_{\text{redll}} = 219/227.$$

– Principal moments of inertia:

$$I_I = 1,348,306/247,168$$

$$I_{II} = 869,596/247,168.$$

– Torsional moment of inertia:

$$I_t = 1,606,083/494,261.$$

– Torsional section modulus:

$$W_t = 7,334/2,177.$$

– Position of principal axes of inertia in relation to road:

$$\alpha = 0^\circ.$$

– Maximum four stress points:

$$S_x = -45/45/45/-45 / 0/35/0/-35,$$

$$S_y = -60/-60/60/60 / -35/0/35/0.$$

The main dimensions according to the side and top views (Figure 7) are (in mm):

$$L_1 = 4,114 \text{ (max.)}$$

$$L_2 = 2,650$$

$$W_1 = 1,517 \text{ (max.)}$$

$$W_2 = 1,147 \text{ (front)}$$

$$W_3 = 1,374 \text{ (rear)}$$

$$H_1 = 1,402 \text{ (max.)}$$

$$H_2 = 1,315$$

$$H_3 = 469 \text{ (box)}$$

Body

The tubular consists of 18 components (Figure 8), e.g.:

- 2 Side pillar,
- 3 Pedal cross-pillar,
- 7 Cockpit cross-pillar,
- 8 Side pillar, front,
- 10 Auxiliary cross-pillar, front,
- 14 Roof frame, side,
- 15 Roof frame, lateral,
- 16 Rear subframe,
- 17 Side pillar, box,
- 19 Bumper, rear.

Support conditions

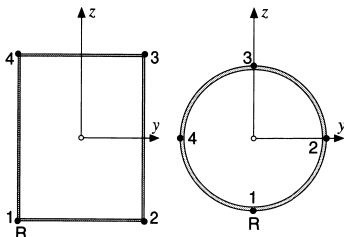
Linear statics, load case – bending (not shown):

A $v_y = 0$; (A – F, see Figure 7).

B, C, E, F $v_z = 0$; D $u_x = 0$, $v_y = 0$.

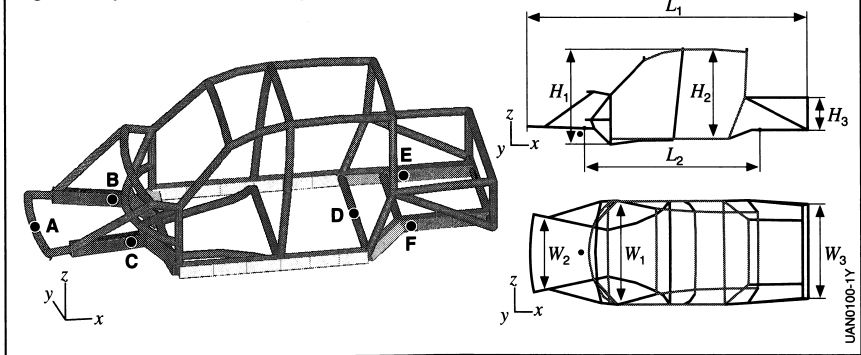
Figure 6: Profile sections of spaceframe

a) Box-type section, b) Tube section.



UAN0099-1Y

Figure 7: Spaceframe of a mini-pickup



Linear statics, load case – torsion:

A, B, C = free;

E, F $v_z = 0$;

D $v_x, v_y = 0, d_y = 0, d_z = 0$.

Free/free dynamics:

No supports, body vibrates freely in the springs.

Loading

Linear statics, load case – torsion:

Torsional moment 3,000,000 Nmm as regular unit load on

Front-axle bearings B: $F = -3593.7$ N

Front-axle bearings C: $F = 3593.7$ N

Free/free dynamics:

Only masses from gravity weight with 170 kg, no supplementary masses.

Results

Free/free dynamics:

- The body can vibrate freely in its springs, the lower eigenvalues and modes of the undamped, elastic system in ascending order; for free/free beginning with several rigid-body vibration modes (here 1 through 6) with the natural frequency 0 in each case.
- The 1st eigenvalue (38 Hz) is obtained as the 1st torsional vibration, the 3rd eigenvalue (48 Hz) as the 1st flexural vibration; with in each case the normalized deformations x, y, z at all nodes (here normalized to max. 0.1 mm), the normalized stresses in all elements, and the deformation processes important for weight optimization per element

and per component 1 through 18 in % as a bar chart (as shown for linear statics in Figure 8). Absolute values are only obtained in the event of an excitation calculation.

Linear statics:

- Deformations x, y, z at all nodes,
- Reaction forces and moments at the support nodes, stresses at all elements, strain energy per element and per component in %.

Weight and stiffness optimization

The computation formula for weight and stiffness optimization has been known since the 1960s (error max. 10 % for doubled rigidity). The stiffness change of the overall structure (in %) is obtained from the stiffness change of the component multiplied by the deformation-process ratio of the component, divided by 100.

This formula is used specifically for optimization purposes in virtually all automobile-critical “Torsion” load cases with consideration of the other load cases. In this way, the structure can be reinforced using pillars that take most of the load, and reduce weight via the pillars that take less of the load.

Application of formula

1st component (see Figure 8):

Component 15 (roof frame, lateral, $G=11.85$ kg with 14.57 % strain energy ratio:

Produces $(14.57 \times 116) / 100 = 16.9$ % change (i.e. reduction) of torsion between the axles when the stiffness (planar moment of inertia) of this component is increased by a factor of 2.169 (116.9 %). The weight increase here is 3.55 kg when the tube diameter is increased from 70 to 90 mm.

2nd component (see Figure 8):

Component 6 (linear reinforcement, $G=11.14$ kg) with 2.64 % strain energy ratio:

Produces $(2.64 \times 250) / 100 = 6.6$ % change (i.e. only very small reduction of torsion between the axles when the stiffness (moment of inertia) of this component is reduced by a factor of 3.5 (250 %). The weight reduction is 3.34 kg when the tube diameter is reduced from 70 to 50 mm.

Result:

A minimal weight increase of $3.55 - 3.34 = 0.21$ kg increases torsional stiffness by $16.9 - 6.6 = 10.3$ % (checking with altered profile sections produces 9 %, this covers the low error in the computation formula).

A look at the strain energy charts of the other load cases (not pictured here) shows that this also applies to torsional vibration and is of no importance to bending. It is thus possible to significantly increase the torsional stiffness and flexural bending stiffness of this body still further and to safely reduce the total weight by reducing the cross-sections of oversized components.

References

- [1] R.W. Clough: The Finite Element Method In Plane Stress Analysis; Publication, 2nd ASCE Conference on Electronic Computation, 1960.
- [2] www.IGFgrothTP2000.de.
- [3] D. Braess: Finite Elemente: Theorie, Schnelle Löser und Anwendungen in der Elastizitätstheorie. 5th Edition, Springer Spectrum, 2013.

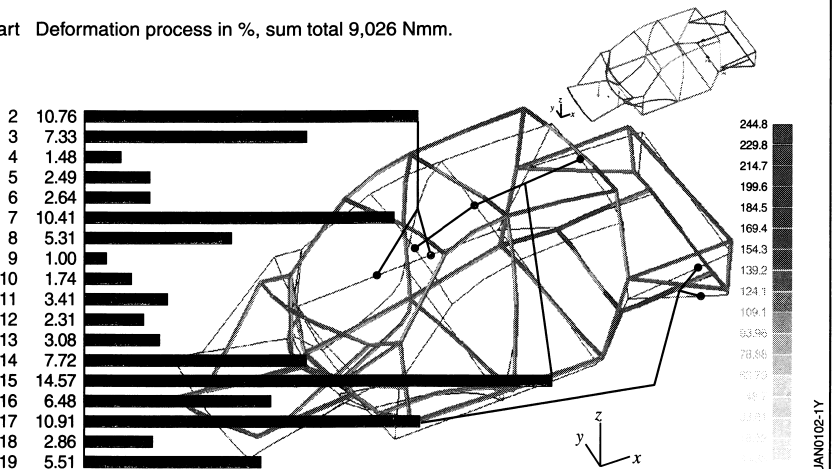
Figure 8: Load case, torsion: 3,000,000 Nmm at the front-axle points

Maximum deformation in mm:

x direction: 0.868459; y direction: 3.071005; z direction: 3.688961.

Maximum stress: 64 N/mm².

Part Deformation process in %, sum total 9,026 Nmm.



Control engineering

Terms and definitions

(in accordance with DIN 19226 [1])

Closed-loop control

In a technical process, the function of closed-loop control is to place and maintain a particular physical parameter (controlled variable y) at a specified value (e.g. the alternator voltage in the electrical system). Here, the controlled variable is continuously measured and compared with the specified value (the reference variable w , e.g. the setpoint input of the alternator voltage as a function of the state of charge of the battery) (Figure 1). In the event of a deviation, a suitable adjustment of the correcting variable u (e.g. influencing the excitation current in the alternator) on the process (controlled system) is carried out in such a way that the controlled

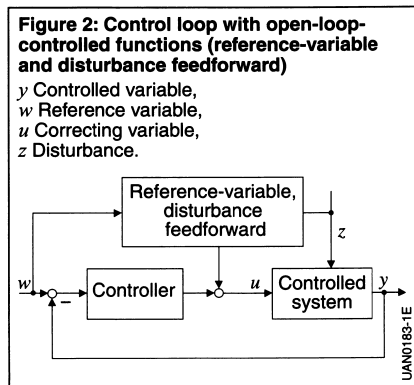
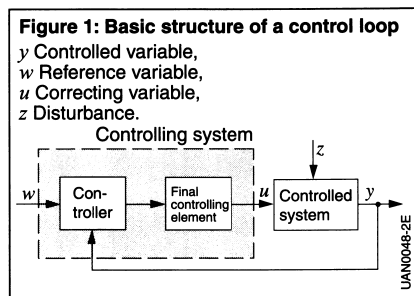
variable y is again adapted to the specification. This process takes place in a closed control loop. Deviations can occur when disturbances z (e.g. the activation of a further electrical consumer) act on the controlled system and affect the controlled variable y in an undesirable way ([2], [3] and [4]).

Closed-loop control operations are performed at many points in a motor vehicle. Further examples include control of cooling-water temperature, A/C control, and many other control operations from the engine (knock control, λ control), transmission (clutch control) and chassis (yaw-rate control).

Open-loop control

Open-loop control is essentially also used instead of closed-loop control. In this case, the closed-loop controlling system is replaced by an open-loop controlling system and feedback of the controlled variable is omitted. This process is only then possible if the behavior of the controlled system is known exactly and if no (non-measurable) disturbances z are acting on the controlled system.

Open-loop control is preferred here as no stability problems can arise on account of there being no feedback. Because the previously mentioned preconditions are rarely satisfied in practice, the use of closed-loop control is for the most part unavoidable.



Combination of open- and closed-loop control

Open- and closed-loop control are, in practice, frequently combined in order to exploit the specific advantages of the two structures. Here, established interconnections between reference variable, disturbance, correcting variable and controlled variable are linked as much as possible in order to realize these as open-loop control. Deviations that still arise on account of changed parameters or non-measurable disturbances are corrected by closed-loop control (Figure 2).

Cascade control

A structure is frequently encountered in which the controlled system is separated into two or more partial systems (e.g. into the actual process and the associated actuator). On the basis of this separation, there are one or more internal controllers and one external controller, which are designed and adjusted separately. This procedure is referred to as cascade control (Figure 3).

The controller design is simplified by this separation of the control task into several manageable subtasks. There are additional advantages in the dynamic response due to the fact that disturbances which act in the internal control loop are corrected there before they affect the external control loop. This makes the entire control operation quicker. Likewise, non-linear characteristic curves of the internal loop can be linearized.

Cascade control is used in many control systems in motor vehicles, e.g. in current control for electrohydraulic actuators or in position control for electric-motor actuators.

Control-engineering transfer elements

A control loop must satisfy four essential requirements in terms of performance:

- The control loop must be stable,
- The control loop must demonstrate a specific stationary accuracy,
- The response to a step change of the reference variable must be sufficiently damped,
- The control loop must be sufficiently fast.

To satisfy these partly contradictory requirements, it is first necessary to describe the static and dynamic responses

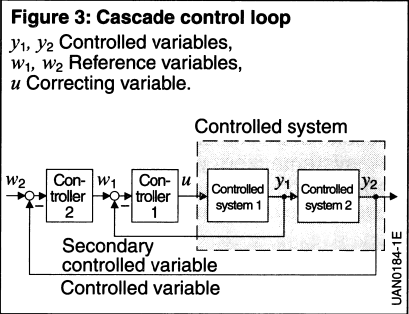


Table 1: Summary of important transfer elements

Transfer element	Differential equation (input u , output y)	Transfer function	Step response
P element	$y = Ku$	K	
I element	$y = K \int_0^t u(\tau) d\tau$	$\frac{K}{s}$	
D element	$y = K \dot{u}$	Ks	
Lag element	$y(t) = Ku(t - T_t)$	$Ke^{-T_t s}$	
$P-T_1$ element	$T \dot{y} + y = Ku$	$\frac{K}{1 + Ts}$	
$P-T_2$ element	$T^2 \ddot{y} + 2dT \dot{y} + y = Ku$	$\frac{K}{1 + 2dT s + T^2 s^2}$	

of the control-loop elements (controlled system and controller) with suitable methods in order to be able to analyze the response in the control loop and to design the controller according to the requirements. This description can be made in the time range (e.g. with differential equations) or in the frequency range (e.g. with a transfer function or a Bode diagram).

Many control-engineering transfer elements can be traced back to specific basic types or can be described by a linking of the same (Table 1).

The task of control-loop synthesis is to design, for a given controlled system, the matching controller (structure and parameters of the transfer element), which fulfills the above-mentioned requirements. There exists for this purpose a series of procedures (e.g. dynamic correction in the Bode diagram, root-locus curve procedures, pole specification, Riccati controller in the state space [4]), which are individually supplemented by specific functions or design steps.

A systematized procedure with the steps described in the following has proven to be useful in practice.

Designing a control task

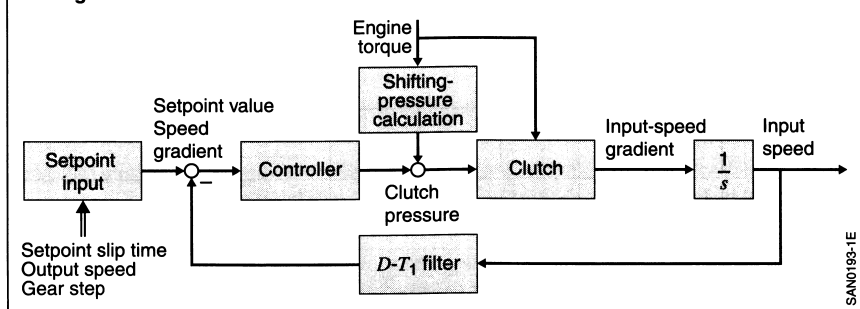
Control task

Typically the control task is not specifically formulated as such, but must be elaborated as the result of requirements in the specific technical process. This involves defining the open- and closed-loop control tasks in the system in order to settle the questions as to what is to be achieved with the control function and with which variables the objective is described. Controlled demand-response shifting in an automatic transmission is mentioned as an example. With this function, the clutch pressure of the shifting clutch is to be aligned during the gear shift to the speed gradient in such a way that the slip time remains constant under all operating conditions, even with variable parameters (e.g. friction coefficient).

System and block diagrams

It is helpful with these considerations, to draw up a system diagram in which the essential interaction between electronics, mechanics, hydraulics and pneumatics and all the sensors, actuators and bus systems can be clearly seen. From this, the control-engineering block diagram must be derived, from which in turn the functional interaction of all the open- and closed-loop-controlled functions with the system to be controlled can be seen (Figure 4). The functions are described in headline form, but not yet formulated in detail.

Figure 4: Control-engineering block diagram, for example of controlled demand-response shifting in an automatic transmission



It should be possible already to obtain a fundamental system understanding of the operative connections from the system diagram and the block diagram. As long as the system (mechanics, peripherals, hardware, etc.) is still in development it should at this point still be possible to exert an influence on the structural design of the system as contemplated by an overall mechatronic approach. The filling behavior of a hydraulically actuated clutch is mentioned as an example. This should be designed in such a way that a reproducible behavior with as little dead time as possible is provided on the basis of line cross-sections, volumes and sealing behavior.

Controlled system

The controlled system is then identified. This can be done theoretically (by modeling) or practically, e.g. by measuring the step response or the frequency response. It is recommended that both methods are followed and an adjustment carried out. Identification is a very comprehensive process, depending on the task. Sometimes it is sufficient to determine only the basic type and the order of the controlled system.

Controller design

The controller is designed on the basis of the result of the identification – this is the central task of controller synthesis. It is advisable for this first to be conducted theoretically using simulations, during which the controller parameters are set at the same time. If this step is sufficiently validated, startup follows on the real controlled system on the test bench or in the vehicle. Recursion steps are usually performed once again here in order to achieve further optimizations.

Design criteria

Over and above this fundamental sequence, reference is made to the following additional criteria:

Digital (discrete-time) control

The majority of control operations in a motor vehicle are implemented on micro-controllers. In this case, it is necessary to suitably establish the sampling time,

oriented towards the dynamic response of the controlled system. It is necessary to ensure here that all the function algorithms can be calculated in the time available between two samplings.

Nonlinearities

In many cases, the simple linear methods described above are not sufficient, since real controlled systems contain nonlinearities (e.g. pressure-regulator characteristics, clutch characteristics, etc.). In simple cases, where static continuous nonlinearities are involved, these can be compensated for by an additional inverse-behavior transfer element. In the case of control operations with small signal amplitudes about an operating point, the system equations can be linearized at this point, otherwise more complex procedures are called for.

Structure switchovers

Many closed-loop control operations are first initiated by open-loop-controlled signals (e.g. first fill clutch, then set shifting pressure, then start shifting control). In these cases, it is necessary when switching from the open-loop- to the closed-loop-controlled state to ensure that this takes place smoothly and that the storage devices (integrators of the I elements) are correctly initialized.

Robustness

The control is typically designed with a “nominal” controlled system. In practice, however, manufacturing tolerances dictate that controlled systems are encountered which deviate up to a defined extent from this nominal definition. Furthermore, the parameters change over the service life, e.g. due to clutch wear or depending on third variables (temperature). In none of these cases should this give rise to significant functional impairments or instability of the control loop. Different procedures from the field of “robust control” or “adaptive control” are available to master these requirements.

Adaptive controllers

Motivation

Controlled systems often do not have a constant behavior. Parameters such as time constants and gains change in many cases. Even the structure of the system can change. Adaptations match open- and closed-loop control processes to a different system behavior. Examples include:

Manufacturing tolerances

Not all the products in a batch are 100 % identical. Because an individual adjustment is complicated, the system should adapt itself automatically to different parameters (e.g. adjustment data for automatic transmissions, [5]).

Wear

Parameters change in response to wear on a reproducible (e.g. increasing clutch control travel) or random level (e.g. friction coefficient of disks). Adaptations can compensate for this different behavior (e.g. clutch-control-travel adaptation for automated clutches, [6]).

Dependence on a third variable (e.g. temperature)

The viscosity of oil is highly temperature-dependent. Because these fluctuations will also occur in the short term (e.g. repeatedly every day), they must be compensated for (e.g. interference-torque observer, lockup-clutch control, [7], [8]).

Dependence on operating point of nonlinear systems

Nonlinear systems are frequently linearized at an operating point and then controlled by a linear controller (small-signal behavior). It is possible by means of an adaptation to take into account the different behavior at the operating points (e.g. adaptation of the shifting pressure for interlaced multiple gearshifts in an automatic transmission, [9]).

The various requirements to master this problem on an open-loop control level result in the desire for adaptive systems, which are described and defined in the following.

Definition of "adaptive control"

The behavior of the control is adapted to the changing properties of the controlled system and its signals. Adaptation procedures are thus essentially divided into two steps.

- Identification: Identification of the system behavior (system parameters) from time-variable values of the system.
- Adaptation: Adaptation of the open- or closed-loop control law as a reaction to a change of the system behavior.

Open-loop-controlled adaptations

Adaptation is performed by open-loop control, a forward-pointing structure. It is assumed here that the changing properties can be recorded by means of measurable external signals z (disturbances)

Figure 5: Open-loop-control adaptive system

y Controlled variable,
 w Reference variable,
 u Correcting variable,
 z Disturbance.

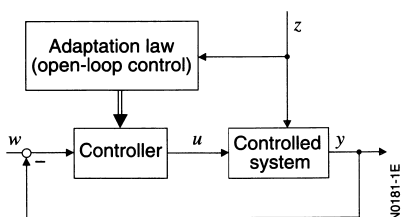
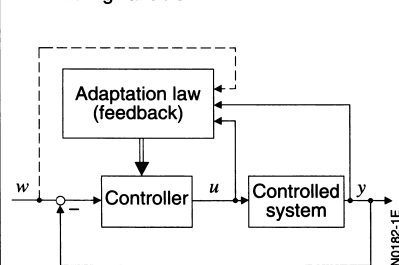


Figure 6: Adaptive system with feedback

y Controlled variable,
 w Reference variable,
 u Correcting variable.



and that the dependence of the control on these signals is known (Figure 5). There is no feedback on internal signals of the control loop to the setting of the controller.

Adaptation with feedback

In the case of an adaptation with feedback, the changing properties cannot be directly recorded and must be identified from measurable signals of the control loop. The identification can be a simple measurement or even a complex algorithm to estimate dynamic process models. In addition to the basic control loop, a second feedback loop is implemented by means of the adaptation law (Figure 6).

Essentially, an open-loop-controlled adaptation is initially recommended, i.e. known and metrologically recordable connections are used in open-loop-controlled form. The advantage of this forward structure can be compared with the advantages of open-loop control over closed-loop control. A feedback loop, which can possibly cause stability problems, is omitted. For the most part, open-loop-control adaptive systems are used in practical industrial applications.

Design notes

The following questions must be clarified when it comes to designing adaptive control:

- Which parameters and features must be adapted because they cannot be covered by a robust design?
- With which signals or variables can these parameters and features be determined?
- Does the system have to be specifically excited in order to determine these parameters and features, or can the adaptation take place during ongoing control operation?
- Can the adaptation be performed by open-loop control means or does feedback have to be provided for?
- How can the stability and convergence of adaptive control be tested?

A detailed treatment of the subjects areas of “Design of adaptive control operations” and “Identification of dynamic processes” can be found in [4], [10] and [11].

References

- [1] DIN 19226: Control technology; terms and designations.
- [2] W. Oppelt: Kleines Handbuch technischer Regelvorgänge. 5th Edition, Verlag Chemie GmbH, 1972.
- [3] O. Föllinger: Regelungstechnik. 12th Edition, VDE Verlag GmbH, 2016.
- [4] R. Isermann: Digitale Regelsysteme. 2nd Edition, Springer-Verlag, 2008.
- [5] DE 102 007 040 485 A1, Abgleichdaten AT-Getriebe.
- [6] DE 102 007 027 702 A1, Kupplungsweg-Adaption für automatisierte Kupplungen.
- [7] DE 000 019943 334 A1, Störmomentbeobachter, Wandlerkupplungsregelung.
- [8] G. Bauer; C. Schwemer: Entwurf einer Wandlerkupplungsregelung unter Berücksichtigung nichtfunktionaler Anforderungen. AUTOREG 2008, 4th Symposium Baden-Baden, 12 and 13 February 2008, VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik.
- [9] DE 102 006 001 899 A1, Adaption Schalldruck bei verschachtelten Mehrfachschaltungen AT-Getriebe.
- [10] R. Isermann: Identifikation dynamischer Systeme 1, 2. 2nd Ed., Springer-Verlag, 2012.
- [11] R. Isermann: Mechatronische Systeme. 2nd Edition, Springer-Verlag, 2007.

Materials

Material parameters

Selecting material for a specific use is based on material parameters that describe the material properties. This process distinguishes between the physical material parameters, which result from the atomic structure, and the mechanical-technological material parameters, which are affected by manufacturing processes, for example.

Some physical material parameters are listed in tables 3 to 6 for a wide variety of materials.

Physical material parameters

Density

Density ρ is the ratio of the mass to the volume of a specific amount of substance (also see DIN 1306:1984-06 [1]). The unit of measurement is kg/m^3 .

Melting temperature

The melting temperature denotes the temperature at which a material transforms from a solid to liquid state.

Boiling temperature

The boiling temperature denotes the temperature at which a material transforms from a liquid into a gas.

Heat of fusion

The specific heat of fusion of a solid (enthalpy of fusion) is the quantity of heat required to transform a material at fusion temperature from the solid to the liquid state. The unit of measurement is kJ/kg .

Heat of evaporation

The specific heat of evaporation of a liquid (enthalpy of evaporation) is the quantity of heat required to evaporate a liquid at constant boiling temperature. It is very pressure-sensitive. The unit of measurement is kJ/kg .

Thermal conductivity

Thermal conductivity is the quantity of heat that flows through a material sample with a defined surface and density due to temperature difference. In the case of liquids and gases (as opposed to solids), thermal conductivity is often highly dependent on temperature. The unit of measurement is $\text{W}/(\text{m}\cdot\text{K})$.

Electric conductivity

Electric conductivity describes the physical property of a material to conduct electric current. According to the Wiedemann-Franz law, it is almost proportional to thermal conductivity. The unit of measurement is S/m (siemens per meter).

Coefficient of thermal expansion

The linear coefficient of thermal expansion α (coefficient of longitudinal expansion) indicates the relative change in length of a material during a temperature variation. For temperature change ΔT , the change in length is defined as $\Delta l = l \alpha \Delta T$. The unit of measurement is K^{-1} .

The cubic coefficient of thermal expansion (coefficient of volume expansion) is defined in a similar manner. The coefficient of volume expansion for gases is roughly $1/273 \text{ K}^{-1}$; For solids, it is roughly three times as large as the coefficient of linear expansion.

Heat capacity

The specific heat capacity (specific heat) is the quantity of heat required to raise the temperature of a substance. It is dependent on temperature.

In the case of gases, we distinguish between specific heat capacity at constant pressure and at constant volume (symbols c_p or c_v). This difference is usually negligible in the case of solid and liquid substances. The unit of measurement is $\text{kJ}/(\text{kg}\cdot\text{K})$.

Permeability

Permeability μ describes the ratio of magnetic flux density B and magnetic field strength H (see magnetic field, relative permeability):

$$B = \mu H.$$

Permeability consists of the magnetic field constant μ_0 and the material-dependent relative permeability μ_r ($\mu_0 = 4\pi \cdot 10^{-7}$ As/Vm, $\mu_r = 1$ for vacuum). Depending on the application in which the magnetic material is used, there are roughly 15 types of permeability. These are defined according to modulation range and type of loading (direct-current or alternating-current field loading). The most important parameters are listed below.

Initial permeability μ_a

The slope of the virgin curve (figure 1, also see hysteresis loop) for $H \rightarrow 0$ is designated as the initial permeability μ_a . In most cases, however, the slope for a specific field strength is specified (in mA/cm) rather than this limit value. Notation: μ_4 is the slope of the virgin curve for $H = 4$ mA/cm.

Maximum permeability $\mu_{\max}$

The maximum slope of the virgin curve is designated as the maximum permeability $\mu_{\max}$.

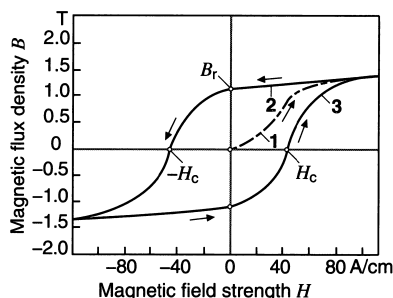
Recoil permeability μ_p

Recoil permeability μ_p (or μ_{rec}) is defined as the average slope of a retrograde magnetic hysteresis loop whose lowest point usually lies on the demagnetization curve:

$$\mu_p = \frac{\Delta B}{\Delta H \mu_0}.$$

Figure 1: Hysteresis loop of a magnetically hard iron grade

1 Rise path,
2, 3 Demagnetization curves.
 H Magnetic field strength,
 B Magnetic flux density,
 B_r Remanence flux density,
 H_c Coercive field strength.
The arrows indicate which magnetic flux density is set if the magnetic field strength is changed.

**Temperature coefficient of magnetic polarization**

The temperature coefficient of the magnetic polarization $TK(J_s)$ indicates the relative change in saturation polarization as temperature changes, it is given in percent per kelvin.

Temperature coefficient of coercive field strength

The temperature coefficient of the coercive field strength $TK(H_c)$ indicates the relative change in coercive field strength as temperature changes, it is given in percent per kelvin.

Curie point

The Curie point (Curie temperature T_C) is the temperature at which the magnetization of ferromagnetic and ferrimagnetic materials becomes zero and at which they behave like paramagnetic materials (see magnetic materials).

Mechanical-technological material parameters

Modulus of elasticity

The modulus of elasticity (E modulus) describes the linear relationship between stress and strain when deforming a solid body in the area of its elastic deformation (image 2). The unit of measurement is MPa. Examples of typical e-moduli are specified in table 1.

Table 1: Modulus of elasticity and Poisson's ratio for some materials

Material	Modulus of elasticity [MPa]	Poisson's ratio
Rubber	100	0.5
Fiber-reinforced plastic (e.g. PA66)	2,000	0.37
Aluminum	70,000	0.34
Titanium	110,000	0.28
Steel	200,000	0.3
Tungsten	400,000	0.28
Ceramics (e.g. Al ₂ O ₃)	400,000	0.23

Poisson's ratio

The dimensionless Poisson's ratio is the proportionality factor between the longitudinal strain of a body under tension or compressive stress and the resulting transverse strain. Typical values are specified in table 1.

Yield point

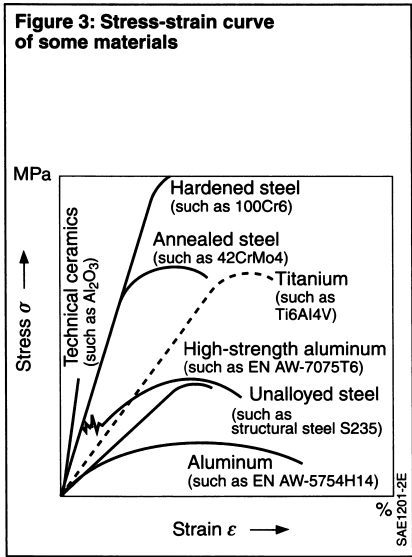
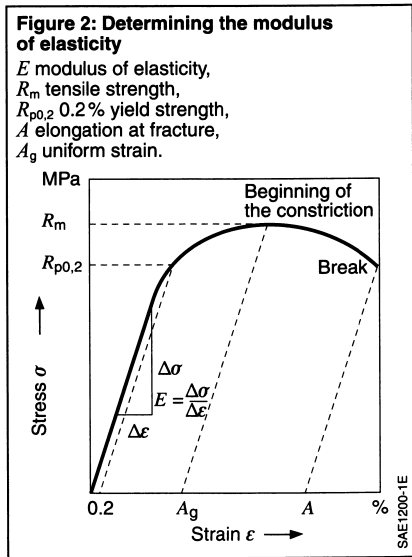
The yield strength (yield point) is the amount of tensile stress starting at which a material experiences lasting plastic deformation. It is determined from the stress-strain curve (σ - ϵ curve) measured in the tensile test (see DIN EN ISO 6892-1:2009-12, [2]). The unit of measurement is MPa.

0.2% yield strength

The 0.2% yield strength ($R_{p0.2}$) is the amount of tensile stress that causes a lasting (plastic) strain of 0.2% in a material. The unit of measurement is MPa.

Tensile strength

The tensile strength R_m is the stress that is calculated in the tensile test from the maximum achieved tractive force relative to the output cross-section of a standardized material sample (figure 2 and figure 3). The unit of measurement is MPa.



Elongation at fracture

The elongation at fracture A is a dimension for the ductility (deformability) of a material. It indicates the lasting extension of a tensile specimen after the fracture relative to the initial length, specified in percent.

Fracture contraction

The fracture contraction Z , like the elongation at fracture, is a dimension for the ductility (deformability) of a material. It indicates in percent the change in cross-sectional area relative to the initial cross-section of a tensile specimen.

Uniform strain

The uniform strain A_g indicates the change in length in percent at which a tensile specimen goes through a plastic deformation without necking. For ductile materials, necking is caused on the specimen (cross-section reduction) after reaching the uniform strain (and tensile strength).

Vibration resistance

Vibration resistance is defined as the deformation and failure behavior of materials under cyclical stress. Typical vibration resistance values such as the fatigue limit have to be determined in often time-con-

suming Wöhler tests (fatigue tests named after August Wöhler) (image 4). These are usually carried out on electromechanical or electrohydraulic testing machines in the frequency range of 10 to 1,000 Hz.

A wide variety of test parameters influence the determined characteristic values (surface quality, test frequency, test duration, test medium, bend on tension or pressure). Here, the determined fatigue limit of a material is well below the structural strength (e.g. R_m). A simple estimate is possible using the FKM guideline (Forschungskuratorium Maschinenbau e.V., see Table 2). The tension-compression fatigue stress factor identifies the ratio of vibration resistance to tensile strength (tension-compression fatigue stress). For metals with a face-centered cubic crystal lattice, such as austenitic steel or aluminum, there is no fatigue limit since the stressability decreases when the stress duration increases.

Knowing the vibration resistance parameters is absolutely necessary for designing parts and components in mechanical engineering, since purely static stress is rare.

Fatigue strength under reversed bending stress

Fatigue strength under reversed bending stress identifies the stress in MPa that a material can withstand without breaking during cyclical bending stress (see vibration resistance).

Figure 4: Depicting the vibration resistance with the Wöhler curve with typical values for the number of stress cycles N

R_m tensile strength,
 $S_{a,D}$ Fatigue limit.

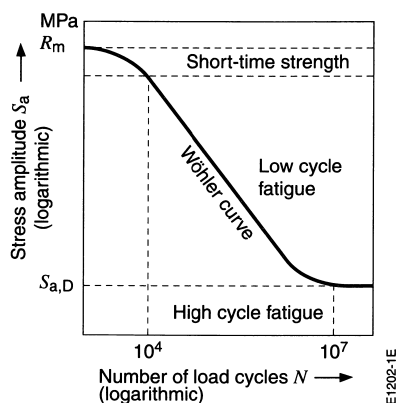


Table 2: Estimation of the vibration resistance

Material group	Tension-compression fatigue stress factor for the number of stress cycles $N = 10^6$
Aluminum, wrought and casting materials	0.30
Cast iron with flake graphite (GJL)	0.30
Cast iron with nodular graphite (GJS)	0.34
Stainless steel	0.40
Structural steel, tempering steel, etc.	0.45

Hardness

The hardness of a material identifies its resistance to the penetration of a body. Various methods for determining hardness are listed in the chapter "Heat treatment of metallic materials".

Fracture toughness

The fracture toughness is the resistance of a material to the spread of cracks. The critical stress intensity factor K_{Ic} is usually used as a characteristic of this. If K_{Ic} is known, the critical fracture load can be determined from crack length, or the critical crack length can be determined from the given external stress value.

Special characteristics for sintered material

Radial crushing strength

The radial crushing strength is a strength parameter which is specified in particular for sintered friction bearings. This is determined by buckling a hollow cylinder during the pressure test.

Porosity

The porosity is a dimensionless characteristic for sintered metals, particularly friction bearings. It describes the portion of part volume that consists of cavities and is not filled with actual material.

Compressive yield point

Comparable to the yield strength in the tensile test, the compressive yield point is added in the pressure test as a material characteristic (unit of measurement MPa). This identifies the stress starting at which an irreversible plastic deformation occurs in the material.

Special characteristics for springs

Bending stress

Bending stress occurs when stress is put on torsion springs. The maximum bending stress here should not exceed $0.7 R_m$.

Shear stress

Shear stress occurs in the material when stress is put on pressure and tension springs. The maximum shear stress (maximum stress) should not exceed $0.5 R_m$. Here, the shear stress is dependent on the coil diameter of the spring, spring force and wire diameter. The difference between maximum stress and minimum stress is also called stress range.

References for material parameters

- [1] DIN 1306:1984-06: Density; concepts, presentation of values.
- [2] DIN EN ISO 6892-1:2017-02: Metallic materials – Tensile testing – Part 1: Method of test at room temperature (ISO 6892-1:2016); German version EN ISO 6892-1:2016.
- [3] D. Radaj, M. Vormwald: Ermüdungsfestigkeit – Grundlagen für Ingenieure [Fatigue strength – Principles for engineers]. 3rd Edition, Springer-Verlag, 2010.
- [4] G. Gottstein: Physikalische Grundlagen der Materialkunde [Physical principles of material science]. 3rd Edition, Springer-Verlag, 2007.
- [5] J. Schijve: Fatigue of Structures and Materials. 2nd Ed., Springer-Verlag, 2009.
- [6] S. Suresh: Fatigue of Materials. 2nd Edition, Cambridge University Press, 1998.

Table 3: Properties of solids

Material/medium		Density	Melting point ¹⁾	Boiling point ¹⁾	Thermal conductivity ²⁾	Mean specific heat capacity ³⁾	Enthalpy of fusion ΔH ⁴⁾	Coefficient of longitudinal expansion ³⁾
		g/cm ³	°C	°C	W/(m · K)	kJ/(kg · K)	kJ/kg	×10 ⁻⁶ /K
Aluminum	Al	2.70	660	2,467	237	0.90	395	23.0
Aluminum alloys		2.60 to 2.85	480 to 655	—	70 to 240	—	—	21 to 24
Amber		1.0 to 1.1	< 300	Decomposes	—	—	—	—
Antimony	Sb	6.69	630.8	1,635	24.3	0.21	172	8.5
Arsenic	As	5.73	—	613 ⁵⁾	50.0	0.34	370	4.7
Asbestos		2.1 to 2.8	< 1,300	—	—	0.81	—	—
Asphalt		1.1 to 1.4	80 to 100	< 300	0.70	0.92	—	—
Barium	Ba	3.50	729	1,637	18.4	0.28	55.8	18.1 to 21.0
Barium chloride		3.86	963	1,560	—	0.38	108	—
Basalt		2.6 to 3.3	—	—	1.67	0.86	—	—
Beef tallow		0.9 to 0.97	40 to 50	< 350	—	0.87	—	—
Beryllium	Be	1.85	1,278	2,970	200	1.88	1,087	11.5
Bismuth	Bi	9.75	271	1,551	8.1	0.13	59	12.1
Bitumen		1.05	< 90	—	0.17	1.78	—	—
Boiler scale		< 2.5	< 1,200	—	0.12 to 2.3	0.80	—	—
Borax		1.72	740	—	—	1.00	—	—
Boron	B	2.34	2,027	3,802	27.0	1.30	2,053	5
Brass CuZn37		8.4	900	1,110	113	0.38	167	18.5
Brickwork		> 1.9	—	—	1.0	0.9	—	—
Bronze CuSn 6		8.8	910	2,300	64	0.37	—	17.5
Cadmium	Cd	8.65	321.1	765	96.8	0.23	54.4	29.8
Calcium	Ca	1.54	839	1,492	200	0.62	233	22
Calcium chloride		2.15	782	> 1,600	—	0.69	—	—
Cellulose acetate		1.3	—	—	0.26	1.47	—	100 to 160
Cement, set		2 to 2.2	—	—	0.9 to 1.2	1.13	—	—
Chalk		1.8 to 2.6	Decomposes into CaO and CO ₂	—	0.92	0.84	—	—
Chamotte (fireclay)		1.7 to 2.4	< 2,000	—	1.4	0.80	—	—
Charcoal		0.3 to 0.5	—	—	0.084	1.0	—	—
Chromium	Cr	7.19	1,875	2,482	93.7	0.45	294	6.2
Chromium oxide	Cr ₂ O ₃	5.21	2,435	4,000	0.42 ⁶⁾	0.75	—	—
Clay, dry		1.5 to 1.8	< 1,600	—	0.9 to 1.3	0.88	—	—
Cobalt	Co	8.9	1,495	2,956	69.1	0.44	268	12.4
Coke		1.6 to 1.9	—	—	0.18	0.83	—	—
Common salt		2.15	802	1,440	—	0.92	—	—
Concrete		1.8 to 2.2	—	—	< 1.0	0.88	—	—
Copper	Cu	8.96	1084.9	2,582	401	0.38	205	—
Cork		0.1 to 0.3	—	—	0.04 to 0.06	1.7 to 2.1	—	—
Corundum, fused		—	—	—	—	—	—	6.5 ⁷⁾
Cotton wadding		0.01	—	—	0.04	—	—	—

¹⁾ At 1.013 bar.²⁾ At 20 °C. ΔH of chemical elements at 27 °C (300 K).³⁾ At 0 to 100 °C.⁴⁾ At melting point and 1.013 bar.⁵⁾ Sublimated.⁶⁾ Powder form.⁷⁾ At 20 to 1,000 °C.

Table 3: Properties of solids (continued)

Material/medium		Density	Melting point ¹⁾	Boiling point ¹⁾	Thermal conductivity ²⁾	Mean specific heat capacity ³⁾	Enthalpy of fusion $\Delta H^4)$	Coefficient of longitudinal expansion ³⁾
		g/cm ³	°C	°C	W/(m · K)	kJ/(kg · K)	kJ/kg	×10 ⁻⁶ /K
Diamond	C	3.5	3,820	—	—	0.52	—	1.1
Duromer								
Phenolic resin without filler		1.3	—	—	0.2	1.47	—	80
Phenolic resin with asbestos fibers		1.8	—	—	0.7	1.25	—	15 to 30
Phenolic resin with wood flour		1.4	—	—	0.35	1.47	—	30 to 50
Phenolic resin with fabric threads		1.4	—	—	0.35	1.47	—	15 to 30
Melamine resin with cellulose fiber		1.5	—	—	0.35	—	—	< 60
Foam rubber		0.06 to 0.25	—	—	0.04 to 0.06	—	—	—
Germanium	Ge	5.32	937	2,830	59.9	0.31	478	5.6
Glass (fused quartz)		—	—	—	—	—	—	0.5
Glass (window glass)		2.4 to 2.7	< 700	—	0.81	0.83	—	< 8
Gold	Au	19.32	1,064	2,967	317	0.13	64.5	14.2
Granite		2.7	—	—	3.49	0.83	—	—
Graphite, pure	C	2.24	< 3,800	< 4,200	168	0.71	—	2.7
Gray cast iron		7.25	1,200	2,500	58	0.50	125	10.5
Hard coal (anthracite)		1.35	—	—	0.24	1.02	—	—
Hard metal K 20		14.8	> 2,000	< 4,000	81.4	0.80	—	5 to 7
Hard rubber		1.2 to 1.5	—	—	0.16	1.42	—	50 to 90 ⁵⁾
Heat-conductive alloy NiCr 8020		8.3	1,400	2,350	14.6	0.50 ⁹⁾	—	—
HR foam, air-filled ⁸⁾		0.015 to 0.06	—	—	0.036 to 0.06	—	—	—
HR foam, freon-filled		0.015 to 0.06	—	—	0.02 to 0.03	—	—	—
Ice (0 °C)		0.92	0	100	2.33 ⁶⁾	2.09 ⁶⁾	333	51 ⁷⁾
Indium	In	7.29	156.6	2,006	81.6	0.24	28.4	33
Iodine	I	4.95	113.5	184	0.45	0.22	120.3	—
Iridium	Ir	22.55	2,447	4,547	147	0.13	137	6.4
Iron, pure	Fe	7.87	1,535	2,887	80.2	0.45	267	12.3
Lead	Pb	11.3	327.5	1,749	35.5	0.13	24.7	29.1
Lead oxide, litharge	PbO	9.3	880	1,480	—	0.22	—	—
Leather, dry		0.86 to 1	—	—	0.14 to 0.16	< 1.5	—	—
Linoleum		1.2	—	—	0.19	—	—	—
Lithium	Li	0.534	180.5	1,317	84.7	3.3	663	56
Magnesium	Mg	1.74	648.8	1,100	156	1.02	372	26.1
Magnesium alloys		< 1.8	< 630	1,500	46 to 139	—	—	24.5
Manganese	Mn	7.47	1,244	2,100	7.82	0.48	362	22
Marble	CaCO ₃	2.6 to 2.8	Decomposes into CaO and CO ₂		2.8	0.84	—	—
Mica		2.6 to 2.9	Decomposes at 700 °C		0.35	0.87	—	3
Molybdenum	Mo	10.22	2,623	5,560	138	0.28	288	5.4
Monel metal		8.8	1,240 to 1,330	—	19.7	0.43	—	—
Mortar, cement		1.6 to 1.8	—	—	1.40	—	—	—
Mortar, lime		1.6 to 1.8	—	—	0.87	—	—	—

¹⁾ At 1.013 bar.²⁾ At 20 °C.³⁾ At 0 to 100 °C.⁴⁾ At melting point and 1.013 bar.⁵⁾ At 20 to 50 °C.⁶⁾ At -20 to 0 °C.⁷⁾ At -20 to -1 °C.⁸⁾ HR foam of phenolic resin, polystyrene, polyethylene or the like, values dependent on cell diameter and filler gas.⁹⁾ At 0 to 100 °C.

Table 3: Properties of solids (continued)

Material/medium		Density	Melting point ¹⁾	Boiling point ¹⁾	Thermal conductivity ²⁾	Mean specific heat capacity ³⁾	Enthalpy of fusion ΔH ⁴⁾	Coefficient of longitudinal expansion ³⁾
		g/cm ³	°C	°C	W/(m·K)	kJ/(kg·K)	kJ/kg	×10 ⁻⁶ /K
Nickel	Ni	8.90	1,455	2,782	90.7	0.46	300	13.3
Nickel silver CuNi12Zn24		8.7	1,020	—	48	0.40	—	18
Niobium	Nb	8.58	2,477	4,540	53.7	0.26	293	7.1
Osmium	Os	22.57	3,045	5,027	87.6	0.13	154	4.3 to 6.8
Palladium	Pd	12.0	1,554	2,927	71.8	0.24	162	11.2
Paper		0.7 to 1.2	—	—	0.14	1.34	—	—
Paraffin		0.9	52	300	0.26	3.27	—	—
Peat dust, air-dried		0.19	—	—	0.081	—	—	—
Phosphorus (white) P		1.82	44.1	280.4	—	0.79	20	—
Pitch		1.25	—	—	0.13	—	—	—
Plaster		2.3	1,200	—	0.45	1.09	—	—
Platinum	Pt	21.45	1,769	3,827	71.6	0.13	101	9
Plutonium	Pu	19.8	640	3,454	6.7	0.14	11	55
Polyamide		1.1	—	—	0.31	—	—	70 to 150
Polycarbonate		1.2	—	—	0.20	1.17	—	60 to 70
Polyethylene		0.94	—	—	0.41	2.1	—	200
Polystyrene		1.05	—	—	0.17	1.3	—	70
Polyvinyl chloride		1.4	—	—	0.16	—	—	70 to 150
Porcelain		2.3 to 2.5	< 1,600	—	1.6 ⁵⁾	1.2 ⁵⁾	—	4 to 5
Potassium	K	0.86	63.65	754	102.4	0.74	61.4	83
Quartz		2.1 to 2.5	1,480	2,230	9.9	0.80	—	8 ⁶⁾ /14.6 ⁷⁾
Radium	Ra	5	700	1,630	18.6	0.12	32	20.2
Red bronze CuSn5ZnPb		8.8	950	2,300	38	0.67	—	—
Red lead, minium Pb ₃ O ₄		8.6 to 9.1	Forms PbO	—	0.70	0.092	—	—
Resin bonded fabric, paper		1.3 to 1.4	—	—	0.23	1.47	—	10 to 25 ⁸⁾
Resistance alloy CuNi 44		8.9	1,280	< 2,400	22.6	0.41	—	15.2
Rhenium	Re	21.02	3,160	5,762	150	0.14	178	8.4
Roofing paper		1.1	—	—	0.19	—	—	—
Rosin		1.08	100 to 130	Decomposes	0.32	1.21	—	—
Rubber, raw		0.92	125	—	0.15	—	—	—
Rubidium	Rb	1.53	38.9	688	58	0.33	26	90
Sand, quartz, dry		1.5 to 1.7	< 1,500	2,230	0.58	0.80	—	—
Sandstone		2 to 2.5	< 1,500	—	2.3	0.71	—	—
Selenium	Se	4.8	217	684.9	2.0	0.34	64.6	37
Silicon	Si	2.33	1,410	2,480	148	0.68	1,410	4.2
Silicon carbide		2.4	Decomposes above 3,000°C	—	9 ⁹⁾	1.05 ⁹⁾	—	4.0
Sillimanite		2.4	1,820	—	151	1.0	—	—
Silver	Ag	10.5	961.9	2,195	429	0.24	104.7	19.2
Slag, blast furnace		2.5 to 3	1,300 to 1,400	—	0.14	0.84	—	—
Sodium	Na	0.97	97.81	883	141	1.24	115	70.6
Soft rubber		1.08	—	—	0.14 to 0.24	—	—	—

¹⁾ At 1.013 bar. ²⁾ At 20°C. ΔH of chemical elements at 27°C (300 K).⁴⁾ At melting temperature and 1.013 bar.⁵⁾ At 0 to 100°C.³⁾ At 0 to 100°C.⁷⁾ Perpendicular to the crystal axis.⁸⁾ At 20 to 50°C.⁶⁾ Parallel to crystal axis.⁹⁾ At 1,000°C.

Table 3: Properties of solids (continued)

Material/medium		Density	Melting point ¹⁾	Boiling point ¹⁾	Thermal conductivity ²⁾	Mean specific heat capacity ³⁾	Enthalpy of fusion ΔH ⁴⁾	Coefficient of longitudinal expansion ³⁾
		g/cm ³	°C	°C	W/(m · K)	kJ/(kg · K)	kJ/kg	×10 ⁻⁶ /K
Soot		1.7 to 1.8	—	—	0.07	0.84	—	—
Steatite		2.6 to 2.7	< 1,520	—	1.6 ⁶⁾	0.83	—	8 to 9 ⁵⁾
Steel, chromium steel		—	—	—	—	—	—	11
Steel, dynamo sheet		—	—	—	—	—	—	12
Steel, high-speed steel		—	—	—	—	—	—	11.5
Steel, magnet steel		—	—	—	—	—	—	11.5
AlNiCo12/6		—	—	—	—	—	—	—
Steel, nickel steel 36 % Ni (invar)		—	—	—	—	—	—	1.5
Steel, sintered		—	—	—	—	—	—	11.5
Steel, stainless (18Cr, 8Ni)		7.9	1,450	—	14	0.51	—	16
Steel, tungsten steel (18 W)		8.7	1,450	—	26	0.42	—	—
Steel, unalloyed and low-alloy		7.9	1,460	2,500	48 to 58	0.49	205	11.5
Sulfur (α)	S	2.07	112.8	444.67	0.27	0.73	38	74
Sulfur (β)	S	1.96	119.0	—	—	—	—	—
Tantalum	Ta	16.65	2,996	5,487	57.5	0.14	174	6.6
Tellurium	Te	6.24	449.5	989.8	2.3	0.20	106	16.7
Thorium	Th	11.72	1,750	4,227	54	0.14	< 83	12.5
Tin (white)	Sn	7.28	231.97	2,270	65.7	0.23	61	21.2
Titanium	Ti	4.51	1,660	3,313	21.9	0.52	437	8.3
Tombac CuZn 20		8.65	1,000	< 1,300	159	0.38	—	—
Tungsten	W	19.25	3,422	5,727	174	0.13	191	4.6
Uranium	U	18.95	1,132.3	3,677	27.6	0.12	65	12.6
Vanadium	V	6.11	1,890	3,000	30.7	0.50	345	8.3
Vulcanized fiber		1.28	—	—	0.21	1.26	—	—
Wax		0.96	60	—	0.084	3.4	—	—
Wood ⁷⁾	Ash	0.72	—	—	0.16	} 2,1 to 2,9	—	} in fiber direction 3 to 4, transverse to fiber 22 to 43
	Balsa	0.20	—	—	0.06		—	
	Beech	0.72	—	—	0.17		—	
	Birch	0.63	—	—	0.14		—	
	Maple	0.62	—	—	0.16		—	
	Oak	0.69	—	—	0.17		—	
	Pine	0.52	—	—	0.14		—	
	Poplar	0.50	—	—	0.12		—	
	Spruce, fir	0.45	—	—	0.14		—	
	Walnut	0.65	—	—	0.15		—	
Wood-wool building slabs		0.36 to 0.57	—	—	0.093	—	—	—
Zinc	Zn	7.14	419.58	907	116	0.38	102	25.0
Zirconium	Zr	6.51	1,852	4,377	22.7	0.28	252	5.8

¹⁾ At 1.013 bar.

²⁾ At 20 °C. ΔH of chemical elements at 27 °C (300 K).

⁴⁾ At melting temperature and 1.013 bar.

⁵⁾ At 20 to 1,000 °C.

⁷⁾ Mean values for air-dried wood (humidity approximately 12%).

Radial thermal conductivity; axial is approx. twice as high.

³⁾ At 0 to 100 °C.

⁶⁾ At 100 to 200 °C.

Table 4: Properties of liquids

Material/medium	Density ²⁾ g/cm ³	Melting point ¹⁾ °C	Boiling point ¹⁾ °C	Thermal conductivity ²⁾ W/(m·K)	Specific heat capacity ²⁾ kJ/(kg·K)	Melting enthalpy $\Delta H^3)$ kJ/kg	Evap- oration enthalpy ⁴⁾ kJ/kg	Coefficient of volume expansion $\times 10^{-3}/K$
Acetone (CH ₃) ₂ CO	0.79	-95	56	0.16	2.21	98.0	523	—
Antifreeze-water mixture 23% by vol.	1.03	-12	101	0.53	3.94	—	—	—
38% by vol.	1.04	-25	103	0.45	3.68	—	—	—
54% by vol.	1.06	-46	105	0.40	3.43	—	—	—
Benzene C ₆ H ₆	0.88	+5.5 ⁶⁾	80	0.15	1.70	127	394	1.25
Common-salt solution 20%	1.15	-18	109	0.58	3.43	—	—	—
Diesel fuel	0.81 to 0.85	-30	150 to 360	0.15	2.05	—	—	—
Ethanol C ₂ H ₅ OH	0.79	-117	78.5	0.17	2.43	109	904	1.1
Ethanol 95% ⁹⁾	0.81	-114	78	0.17	2.43	—	—	—
Ethyl chloride C ₂ H ₅ Cl	0.90	-136	12	0.11 ⁵⁾	1.54 ⁵⁾	69.0	437	—
Ethyl ether (C ₂ H ₅) ₂ O	0.71	-116	34.5	0.13	2.28	98.1	377	1.6
Ethylene glycol C ₂ H ₄ (OH) ₂	1.11	-12	198	0.25	2.40	—	—	—
Fuel oil EL	< 0.83	-10	> 175	0.14	2.07	—	—	—
Gasoline/petrol	0.72 to 0.75	-50 to -30	25 to 210	0.13	2.02	—	—	1.0
Glycerin C ₃ H ₅ (OH) ₃	1.26	+20	290	0.29	2.37	200	828	0.5
Hydrochloric acid 10 % HCl	1.05	-14	102	0.50	3.14	—	—	—
Kerosene	0.76 to 0.86	-70	> 150	0.13	2.16	—	—	1.0
Linseed oil	0.93	-15	316	0.17	1.88	—	—	—
Lubricating oil	0.91	-20	> 300	0.13	2.09	—	—	—
Mercury ⁸⁾	13.55	-38.84	356.6	10	0.14	11.6	295	0.18
Methanol CH ₃ OH	0.79	-98	65	0.20	2.51	99.2	1,109	—
Methyl chloride CH ₃ Cl	0.99 ⁷⁾	-92	-24	0.16	1.38	—	406	—
m-xylene C ₆ H ₄ (CH ₃) ₂	0.86	-48	139	—	—	—	339	—
Nitric acid, conc. HNO ₃	1.51	-41	84	0.26	1.72	—	—	—
Paraffin oil	—	—	—	—	—	—	—	0.764
Petroleum ether	0.66	-160	> 40	0.14	1.76	—	—	—
Rapeseed oil	0.91	± 0	300	0.17	1.97	—	—	—
Silicone oil	0.76 to 0.98	—	—	0.13	1.09	—	—	—
Sulfuric acid, conc. H ₂ SO ₄	1.83	+10.5 ⁶⁾	338	0.47	1.42	—	—	0.55

Table 4: Properties of liquids (continued)

Material/medium	Density ²⁾ g/cm ³	Melting point ¹⁾ °C	Boiling point ¹⁾ °C	Thermal conductivity ²⁾ W/(m·K)	Specific heat capacity ²⁾ kJ/(kg·K)	Melting enthalpy $\Delta H^3)$ kJ/kg	Evap- oration enthalpy ⁴⁾ kJ/kg	Coefficient of volume expansion $\times 10^{-3}/K$
Tar	1.2	-15	300	0.19	1.56	—	—	—
Toluene C_7H_8	0.87	-93	111	0.14	1.67	74.4	364	—
Transformer oil	0.88	-30	170	0.13	1.88	—	—	—
Trichloroethylene C_2HCl_3	1.46	-85	87	0.12	0.93	—	265	1.19
Turpentine oil	0.86	-10	160	0.11	1.80	—	293	1.0
Water	1.00 ¹⁰⁾	± 0	100	0.60	4.18	332	2,256	0.18 ¹¹⁾

¹⁾ At 1.013 bar. ²⁾ At 20 °C. ³⁾ At melting temperature and 1.013 bar. ⁴⁾ At boiling temperature and 1.013 bar. ⁵⁾ At 0 °C.

⁶⁾ Setting temperature 0 °C. ⁷⁾ At -24 °C. ⁸⁾ For conversion from Torr to Pa, use 13.5951 g/cm³ (at 0 °C).

⁹⁾ Denatured ethanol. ¹⁰⁾ At 4 °C. ¹¹⁾ Volume expansion on freezing: 9%.

Table 5: Properties of water vapor

Absolute pressure bar	Boiling point °C	Evaporation enthalpy kJ/kg
0.1233	50	2,382
0.3855	75	2,321
1.0133	100	2,256
2.3216	125	2,187
4.760	150	2,113
8.925	175	2,031
15.55	200	1,941
25.5	225	1,837
39.78	250	1,716
59.49	275	1,573
85.92	300	1,403
120.5	325	1,189
165.4	350	892
221.1	374.2	0

Table 6: Properties of gases

Material/medium		Density ¹⁾	Melting point ²⁾	Boiling point ²⁾	Thermal conductivity ³⁾	Specific heat capacity ³⁾			Enthalpy of evaporation ²⁾
		kg/m ³	°C	°C	W/(m·K)	c _p	c _v		kJ/kg
Acetylene	C ₂ H ₂	1.17	−84	−81	0.021	1.64	1.33	1.23	751
Air		1.293	−220	−191	0.026	1.005	0.716	1.40	209
Ammonia	NH ₃	0.77	−78	−33	0.024	2.06	1.56	1.32	1,369
Argon	Ar	1.78	−189	−186	0.018	0.52	0.31	1.67	163
n-butane	C ₄ H ₁₀	2.70	−138	−0.5	0.016	1.67	1.51	1.10	—
i-butane	C ₄ H ₁₀	2.67	−145	−10.2	0.016	—	—	1.11	—
Blast-furnace gas		1.28	−210	−170	0.024	1.05	0.75	1.40	—
Carbon disulfide	CS ₂	3.41	−112	−46	0.0073	0.67	0.56	1.19	—
Carbon dioxide	CO ₂	1.98	−57 ⁴⁾	−78	0.016	0.82	0.63	1.30	368
Carbon monoxide	CO	1.25	−199	−191	0.025	1.05	0.75	1.40	—
Chlorine	Cl ₂	3.21	−101	−35	0.009	0.48	0.37	1.30	288
City/town gas		0.56 to 0.61	−230	−2010	0.064	2.14	1.59	1.35	—
Cyanogen (dicyan)	(CN) ₂	2.33	−34	−21	—	1.72	1.35	1.27	—
Dichlorodifluoromethane (= Freon F 12)	CCl ₂ F ₂	5.51	−140	−30	0.010	0.61	0.54	1.14	—
Ethane	C ₂ H ₆	1.36	−183	−89	0.021	1.66	1.36	1.22	522
Ethanol vapor	C ₂ H ₄	2.04	−114	+78	0.015	—	—	1.13	—
Ethylene		1.26	−169	−104	0.020	1.47	1.18	1.24	516
Fluorine	F ₂	1.70	−220	−188	0.025	0.82	0.61	1.35	172
Helium	He	0.18	−270	−269	0.15	5.20	3.15	1.65	20
Hydrogen	H ₂	0.09	−258	−253	0.181	14.39	10.10	1.42	228
Hydrogen chloride	HCl	1.64	−114	−85	0.014	0.81	0.57	1.42	—
Hydrogen sulfide	H ₂ S	1.54	−86	−61	0.013 ¹⁾	0.96	0.72	1.34	535
Krypton	Kr	3.73	−157	−153	0.0095	0.25	0.15	1.67	108
Methane	CH ₄	0.72	−183	−164	0.033	2.19	1.68	1.30	557
Methyl chloride	CH ₃ Cl	2.31	−92	−24	—	0.74	0.57	1.29	446
Neon	Ne	0.90	−249	−246	0.049	1.03	0.62	1.67	86
Nitrogen	N ₂	1.24	−210	−196	0.026	1.04	0.74	1.40	199
Oxygen	O ₂	1.43	−218	−183	0.0267	0.92	0.65	1.41	213
Ozone	O ₃	2.14	−251	−112	0.019	0.81	0.63	1.29	—
Propane	C ₃ H ₈	2.00	−182	−42	0.018	1.70	1.50	1.13	—
Propylene	C ₃ H ₆	1.91	−185	−47	0.017	1.47	1.28	1.15	468
Sulfur dioxide	SO ₂	2.93	−73	−10	0.010	0.64	0.46	1.40	402
Sulfur hexafluoride	SF ₆	6.16 ³⁾	−50.8	−63.9	0.011	0.66	—	—	117 ¹⁾
Water vapor, 100 °C ⁵⁾		0.60	± 0	+100	0.025	2.01	1.52	1.32	—
Xenon	Xe	5.89	−112	−108	0.0057	0.16	0.096	1.67	96

¹⁾ At 0 °C and 1.013 bar.²⁾ At 1.013 bar.³⁾ At 20 °C and 1.013 bar.⁴⁾ At 5.3 bar.⁵⁾ At saturation and 1.013 bar, also see Table 5.

Material groups

Material requirements

Technically and economically beneficial material use requires the knowledge of the application and the stress occurring in operation to derive the material requirements. The goal is to implement the desired function in the part or component and guarantee it throughout the planned period of use. The crucial loads here are usually of mechanical nature (strength, modulus of elasticity, see Table 1), but are often supplemented by corrosion, wear or temperature loads. Furthermore, specific physical properties (e.g. magnetism, conductivity) or a combination of those mentioned can be critical for material use.

Classification of materials

Various options exist for classifying the materials now used in technical applications. One example is subdividing into the four material groups:

- Metallic materials: smelted metals, sintered metals.
- Non-metallic inorganic materials: ceramics, glass.
- Non-metallic organic materials: natural substances, plastics.
- Composite materials.

Table 1: E-modulus of some materials (example)

Material group	Material type (selection)	Abbreviation Example	E-modulus [GPa]
Cast iron and malleable cast iron	Cast iron with lamellar graphite	EN-GJL-200	80 to 140
	Cast iron with spheroid graphite	EN-GJS-400	160 to 180
	White malleable cast iron	EN-GJMW-450	175 to 195
	Black malleable cast iron	EN-GJMB-450	175 to 195
Cast steel	Cast steel for general use	GE300	≈ 210
Steel	Unalloyed steels	C60	≈ 210
	Low-alloy steels	42CrMoS4	≈ 210
	Austenitic steels	X5CrNi18-10	≥ 190
	High-alloy tool steels	HS 6-5-4	≤ 230
Wrought copper alloys	High-conductivity copper	Cu-ETP	110 to 130
	Brass	CuZn30	115
	Nickel silver	CuNi18Zn20	135
	Tin bronze	CuSn6	100 to 120
Casting copper alloys	Tin cast bronze	CuSn10-C	100
	Red brass	CuSn7Zn4Pb7-C	100
Aluminum alloys	Wrought aluminum alloy	EN AW-AISi1MgMn	65 to 75
	Casting aluminum alloy	EN AC-AISi12Cu1	65 to 80
Magnesium alloys	Wrought magnesium alloy	MgAl6Zn	40 to 45
	Cast magnesium alloy	EN-MCMgAl9Zn1	40 to 45
Titanium alloys	Wrought titanium alloy	TiAl6V4	110
Tin alloys	Cast tin alloys for composite slide bearings	SnSb12Cu6Pb	30
Zinc alloys	Die-cast zinc alloy	GD-ZnAl4Cu1	70

Metallic materials

The metals generally have a crystalline structure. Their atoms are arranged in a regular crystal lattice. The valence electrons of the atoms are not bound to a specific atom, but are able to move freely within the metal lattice – a metallic bond is present. This special metal-lattice structure explains the characteristic properties of metals: the ductility and resulting high level of formability, the high electric conductivity, the high level of thermal conductivity, the low light transmission and the great optical reflective ability (metallic gloss).

The metallic materials used most often in technical applications include ferrous material. Iron base alloys are primarily used for this. An alloy is defined as a metallic material that consists of at least two chemical elements.

The ferrous materials are divided into groups for steels and cast iron materials. The main difference between the two groups is carbon content. The carbon content of steel is generally less than 2% and over 2% for cast iron (for details, see EN metallurgy standards).

Steels

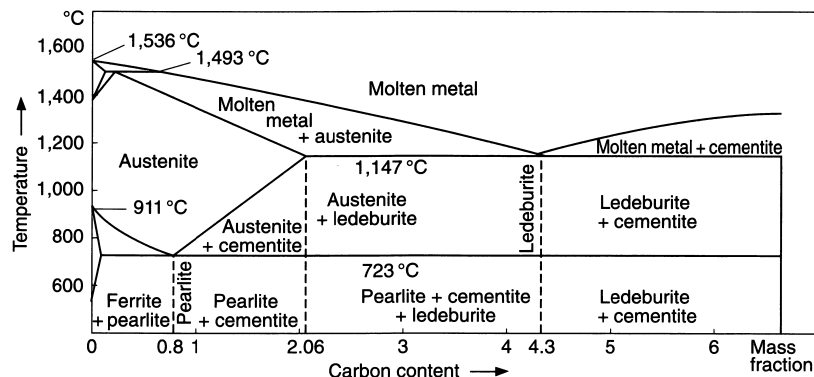
Structural compositions of steels

The majority of steels primarily consists of iron and other property-determining alloying elements such as chrome, nickel, vanadium, molybdenum and titanium. The most important alloying element is carbon, which is diffused in the iron lattice and present as iron carbide (Fe_3C), also called cementite. Various phases are present in the steel with their characteristic properties depending on the carbon content, the alloying elements and heat treatment (e.g. hardening, quenching and drawing). They can be found in the iron/carbon diagram (Figure 1). The most important phases are described below.

Ferrite

The ferritic structure exhibits a body-centered cubic atomic arrangement. It is soft, easily formable and ferromagnetic (e.g. S235). In unalloyed steels, ferrite is only present at very low carbon contents (<0.02%) as an individual phase (also called α phase).

Figure 1: Iron/carbon diagram



Pearlite

Pearlite has a lamellar iron-carbon structure, consisting of alternating associated areas of iron carbide (Fe_3C) and ferrite. This phase is present in steels with a carbon content higher than 0.02%. As carbon content increases, the pearlite content present increases along with the ferrite until the structure consists entirely of pearlite at a carbon content of 0.8% (e.g. C80). The pearlitic structure exhibits high strength.

Increasing the carbon content further results in the increased formation of iron carbide, which tends to separate out at the grain boundaries.

Austenite

If a steel is heated (between approximately 750°C and 1,150°C depending on the carbon content), the body-centered cubic iron lattice (α phase) transforms into a face-centered cubic iron lattice (γ phase). This austenitic structure can diffuse 100 times as much carbon as the ferritic structure. Ferrite and pearlite then form from the structure during the long cooling periods. In simple iron-carbon alloys, this lattice structure only exists at these high temperatures. Adding alloying elements such as nickel or cobalt allows the austenitic structure to be preserved stably even at room temperature. These types of austenitic alloys are very tough, easily formable, and resistant to heat and corrosion (e.g. X5CrNi1810).

Martensite

When cooling a steel very quickly from the austenite area during heat treatment (e.g. by quenching in water or oil), the carbon diffused in the lattice does not have enough time to transform into ferrite and pearlite. The austenitic structure abruptly deforms into a distorted lattice structure called martensite. The martensitic structure is very hard and brittle (e.g. 100Cr6).

Steel types

Along with many different organization and designation options, steels are often named based on their typical application, use or machining technique. Table 2 and Table 3 give an overview of properties and typical applications.

Structural steel

This is a simple steel that can usually be welded or machined easily. It is used in large quantities in general mechanical engineering and civil engineering and is reasonably priced (e.g. S235JR).

Free-cutting steel

This steel typically features high sulfur content, which causes short chips to form when machining (turning, milling). These chips can be extracted in the machinery without complications. These types of steels are usually used for volume components with low to medium-strength stresses (e.g. 11SMn30).

Tool steel

This is the umbrella term for a type of steel used for tools and molds. Depending on the application, it can feature high tensile strength, thermal stability, viscosity or corrosion resistance. This steel type can be classified as a cold-work steel (up to approx. 200°C) or hot-work steel (up to approx. 400°C), depending on the operating temperature. High-speed steels with high thermal stability (greater than 400°C) and resistance to wear are also used (e.g. 90MnCrV8, X40CrMoV5-1, HS 6-5-2), particularly for machining tools such as drills and milling cutters.

Tempering steel

This type of steel is particularly well-suited for quench and draw heat treatment (hardening and draw tempering). This steel is often used for dynamically stressed shafts due to the high tensile strength and outstanding viscosity (e.g. 42CrMo4).

Table 2: Properties and typical applications for steels

Material group	Example materials	Surface hardness	Core hardness	Properties	Typical applications
Case-hardening steel (case-hardened and tempered)	C15	700 to 850 HV	200 to 450 HV	High resistance to wear	Moderately stressed gear wheels, joints
	16MnCr5	700 to 850 HV	300 to 450 HV	High resistance to wear	Piston pins, gear wheels
	18CrNiMo7-6	700 to 850 HV	400 to 550 HV	High resistance to wear, weldable	Gear wheels, drive parts
Nitriding steel (quenched, tempered and nitrided)	31CrMoV9	700 to 850 HV	250 to 400 HV	High resistance to wear	Highly-stressed piston pins, spindles
Ball and roller bearing steel (hardened and tempered)	100Cr6	60 to 64 HRC		High resistance to wear, high hardness and strength	Ball bearings, rolling elements, needle bearings
Cold-work steel (unalloyed, hardened and tempered)	C80	60 to 64 HRC		Surface hardened, tough core	Blades, chisels, striking tools
Cold-work steel (alloyed, hardened and tempered)	90MnCrV8	60 to 64 HRC		High dimensional stability, great edge retention and toughness.	Reamers, woodworking tools, thread cutters
	X210Cr12	60 to 64 HRC		Highly wear-resistant, great dimensional stability, high hardening properties	Cutting and punching tools, broaching tools, flange rollers, cold extrusion tools
Hot-work steel (hardened and tempered)	X40Cr-MoV5-1	43 to 45 HRC		High resistance to heat-related wear, great toughness and thermal conductivity	Forging dies, die-casting tools, extrusion dies
High-speed steel (hardened and tempered)	HS 6-5-2	61 to 65 HRC		High strength, great toughness, high hardness	Drills, countersinking tools
Stainless martensitic steel (hardened and tempered)	X20Cr13	≈40 HRC		Usable up to 550 °C, ferromagnetic, high resistance to wear	Pump parts, piston rods, nozzle needles, ship propellers
	X90Cr-MoV18	≥57 HRC		Good chemical resistance, resistance to wear and polishing properties	Cutting tools, ball bearings

Table 3: Properties and typical applications for steels

Material group	Example materials	Tensile strength [MPa]	Yield point [MPa]	Elongation at fracture A5 [%]	Fatigue strength under reversed bending stress [MPa]	Properties	Typical application examples
Hot-galvanized strip/sheet	DX53D	≤ 380	≤ 260	≥ 30 (A80)	≥ 190	Good deep-drawing capability, hot-galvanized, higher strength	Automotive sheet metal parts e.g. engine hood, side panel
Structural steel	S235 JR	340 to 510	≥ 225	≥ 26	≥ 170	Good weldability, no preceding or retroactive heat handling	General construction parts, profiles
Free-cutting steel	11SMn30	380 to 570	–	–	≥ 190	High sulfur content for short chips, not suitable for welding	Moderately stressed volume components from the automotive industry such as shafts
	35S20	520 to 680	–	–	≥ 260	High sulfur content for short chips, not suitable for welding	Volume parts such as moderately stressed screws
Tempering steel	C45	700 to 850	≥ 490	≥ 14	≥ 280	Tough, heat treatable, moderate to high strength	Axles, bolts, screws
	42CrMo4	1,100 to 1,300	≥ 900	≥ 10	≥ 440		Crankshafts, axle shafts
	30CrNiMo8	1,250 to 1,450	≥ 1,050	≥ 9	≥ 500		Drive parts
Stainless ferritic steel (annealed)	X6Cr17	450 to 600	≥ 270	≥ 20	≥ 200	Good corrosion resistance, deep-draw capability, polishing properties, magnetic	Household devices, medical technology, sanitary
Stainless austenitic steel (solution annealed)	X5CrNi18-10	500 to 700	≥ 190	≥ 45	–	Most frequently used corrosion resistant steel, good weldability, good deep-draw and polishing capability, non-magnetic	Chemical apparatus engineering, architecture, household objects
Cold strip (unalloyed)	DC05LC	270 to 330	≤ 180	≥ 40 (A80)	≥ 130	Good deep-draw properties	Automotive sheet metal parts e.g. engine hood, side panel

Table 4: Properties and typical applications for spring steel

Material	Diameter and thickness [mm]	Modulus of elasticity E [MPa]	Shear modulus G [MPa]	Minimum tensile strength [MPa]	Fracture contraction Z [%]	Permissible bending stress [MPa]	Permissible stress range for $N \geq 10^7$ [MPa]	Permissible maximum stress [MPa]	Properties and applications
Spring steel wire D, (patented and drawn cold-hammered)	1	206,000	81,500	2,230	40	1,590	380	1,115	Tension, compression, torsion springs, high static and low dynamic stress capacity
	3			1,840	40	1,280	360	920	
	10			1,350	30	930	320	675	
Stainless spring steel wire	1	185,000	73,500	2,000	40	1,400	–	1,000	Stainless springs
	3			1,600	40	1,130	–	800	
Heat-treated, alloyed valve-spring steel wire VD SiCr	1	200,000	79,000	2,060	50	–	430	1,030	High-strength spring material, high toughness e.g. passenger car valve springs
	3			1,920	50	–	430	960	
Spring steel strip Ck85	≤ 2.5	206,000	–	1,470	–	1,270	640	–	Highly stressed leaf springs
Stainless spring steel strip	≤ 1	185,000	–	1,370	–	1,230	590	–	Stainless leaf springs

Case-hardening steel

These steels feature a low carbon content (e.g. C15), which is usually increased up to 0.8% by inserting the steel components into a carbonaceous furnace atmosphere (approx. 900 °C) in the edge region. The subsequent quenching causes martensite to form in the edge region and, in conjunction, high surface hardening at a simultaneously viscous core (ferritic, pearlitic). Applications include gear wheels and piston pins. Unlike nitriding, case hardening causes greater delays due to the martensitic hardening.

Nitriding steel

This steel forms a very hard and wear-resistant surface layer (approx. 10 µm) due to its chemical composition in combination with a corresponding heat treatment (nitriding at approx. 450 to 600 °C). Here, the viscous material core remains preserved, which makes this material particularly well-suited for gear wheels. There is less component deformation due to the lower temperature in comparison to case hardening (e.g. 31CrMoV9).

Spring steel

These are steels that have a particularly fine structure with high yield strength due to alloying silicon. They are especially well-suited for manufacturing springs in engine production (e.g. valve springs). Table 4 gives an overview of various spring steels.

Cast iron materials

Cast iron is defined as an iron-based alloy with a carbon content greater than approximately 2%. The cast iron designation derives from these alloys' particularly high level of castability. One reason for this is the significantly lower melting temperature in comparison to steel (approximately 1,150 °C compared to approximately 1,500 °C).

Table 5 gives an overview of typical properties and applications for some materials.

Table 5: Properties and typical applications for cast irons

Material group	Example materials	Tensile strength [MPa]	Yield point [MPa]	Elongation at fracture A5 [%]	Fatigue strength under reversed bending stress [MPa]	Properties	Typical application examples
Cast iron with lamellar graphite (gray cast iron)	EN-GJL-200	≥ 200	–	–	90	Good castability, high vibration damping and heat capacity	Pump housing, cylinder head, machine bed, brake disc
Cast iron with spheroid graphite	EN-GJS-400-15	≥ 400	≥ 250	≥ 15	200	Higher strength and ductility than GJL	Crankshaft, crank-case, hubs
White malleable cast iron	EN-GJMW-400-5	≥ 400	≥ 220	≥ 5 (A3)	–	Good weldability, thin wall thicknesses, castable	Fittings
Black malleable cast iron	EN-GJMB-350-10	≥ 350	≥ 200	≥ 10 (A3)	–	Better machinability and hardening capability, can be manufactured with thick walls as white malleable cast iron	Fittings

Types of cast iron

The slow cooling during casting results in carbon mostly being present as graphite, typically in a ferritic-pearlitic matrix.

Gray cast iron

Depending on the formation of the graphite in layers or spheroids, the cast iron is referred to as either gray cast iron with lamellar graphite (GJL) or gray cast iron with spheroid graphite (GJS). The spheroid formation of the graphite is controlled in this process by treating the molten metal accordingly, such as with calcium. GJS features higher strength and better formability than GJL. Gray cast iron with vermicular graphite represents the middle ground between these two alloy types, both in terms of both graphite formation as well as properties.

Malleable cast iron

Along with these three types, the structure can also be affected by heat treatment (malleablizing) after casting. For malleable cast iron, the casted component made of white cast iron is subjected to an annealing process, where the iron carbide present in the matrix is converted into temper carbon. This increases the viscosity of the material. Based on the appearance of the fracture surface, these are referred to as white (annealed in an oxidizing, decarburizing atmosphere) and black malleable cast irons (annealed in a neutral atmosphere).

Cast steel

Cast steel is used for particularly highly stressed, weldable castings. In comparison to simple cast iron with lamellar graphite, castability is significantly worse while casting temperature and shrinkage are considerably higher.

Nonferrous metals

Along with the steels, some nonferrous metals (NF metals) have important technically relevant properties, such as those required in various applications. Special emphasis should be placed on copper alloys (heavy metal, density greater than 5 g/cm^3), which are used in many applications due to their high electric conductivity, and easily machinable, lightweight and nonmagnetic aluminum materials (light alloy, density less than 5 g/cm^3).

Classification of NF metals

NF metals (nonferrous metals and alloys with an iron percentage of less than 50%) can generally be classified into wrought and cast alloys. Here, NF wrought alloys exhibit a very well malleable, ductile structure and can be brought to the desired shape by deep-drawing, bending, flanging, extruding or using similar deformation processes. Due to low ductility, shaping is limited to casting for NF casting alloys.

The properties of NF metals are frequently adapted to material requirements by adding alloying elements. Tables 6 through 9 give an overview of typical properties and applications of some materials.

Aluminum alloys

In aluminum alloys, adding alloying elements primarily aims to improve mechanical properties. For example, adding zinc or copper makes heat treatment possible, which results in high strength.

Copper alloys

Alloying zinc to copper forms brass (e.g. CuZn30), which leads to a higher strength compared to pure copper. Copper alloys with other alloying elements other than zinc are called bronzes. The most common bronzes are tin bronzes (e.g. CuSn6) with high strength and low ductility. Copper-nickel alloys (nickel-silver) are frequently used for plug contacts.

Table 6: Properties and typical applications for wrought copper alloys

Material group	Typical materials	Tensile strength [MPa]	Yield strength [MPa]	Fatigue strength under reversed bending stress (reference values) [MPa]	Properties	Application examples
High-conductivity copper	Cu-OF R220	220 to 260	≤140	70	High electric conductivity, solderable, weldable, deformable	Electronics, electrical engineering
Brass	CuZn30 R350	350 to 430	>170	110	Good strength properties, excellent cold-workability, can be soldered and brazed very well	Deep-drawn parts, fastening elements, screws, zippers
Nickel silver	CuNi18Zn20 R500	500 to 590	> 410	–	Cold-workable, non-tarnishing, good spring properties	Relay springs, plugs
Tin bronze	CuSn6 R470	> 470	≈ 350	190	Good cold-workability, good strength properties, good hardening properties, good wear and corrosion resistance, solderable	Springs, metal hoses, general machine and apparatus engineering

Table 7: Properties and typical applications for NF wrought alloys

Material group	Typical materials	Tensile strength [MPa]	Yield strength [MPa]	Fatigue strength under reversed bending stress (reference values) [MPa]	Properties	Application examples
Pure aluminum	EN AW-Al99,5 O	65	20	40	High conductivity, ductile, food-grade	Apparatus and tank engineering, deep-drawn parts
Self-hardening aluminum alloy	EN AW-AlMg2Mn0,8 H111	190	80	90	Seawater resistant, easily weldable	Automotive and ship engineering
Heat-treatable aluminum alloy	EN AW-AlSi1MgMn T6	310	260	90	Artificially aged, good combination of strength, corrosion resistance	Most frequently used heat-treatable aluminum alloy, profiles, vehicle frames, automobile transverse links
Heat-resistant aluminum alloy	EN AW-AlCu4MgSi (A) T4	390	245	120	Very good machinability, high temperature stability, good thermal stability	Components in hydraulic systems, aviation industry
High-strength aluminum alloy	EN AW-AlZn5,5MgCu T6	540	485	140	Good machinability, extremely high strength	Aviation industry, mechanical engineering
Universal aluminum casting alloy	EN AC-AlSi7Mg0,3 T6	290	210	80	Universal casting alloy, good mechanical properties, good corrosion resistance, very good weldability and machinability	Fittings, motor production, architecture
Eutectic casting alloy	EN AC-AlSi12Cu1 (Fe) DF	240	140	70	Excellent mold filling capability and casting characteristics, good chemical resistance	Thin-walled castings, fittings
Universal aluminum die-cast alloy	EN AC-AlSi9Cu3 (Fe) DF	240	140	70	Good processibility, low cost	Intake manifolds, transmission housings

Table 8: Properties and typical applications for NF casting alloys

Material group	Typical materials	Tensile strength	Yield strength	Fatigue strength under reversed bending stress (reference values)	Properties	Application examples
		[MPa]	[MPa]	[MPa]		
High-strength casting alloy	EN AC-AICu4Ti K T6	330	220	90	High strength and toughness	Transmission housings
Magnesium die-cast alloy	EN-MCMgAl9Zn1	200	140	50	Very light, low corrosion and thermal resistance, chips are combustible	Covers, housing, mobile telephone parts
Titanium forge alloy	TiAl6V4	890	820	–	High strength and corrosion resistance	Implants, aviation industry

Other NF metals

Along with aluminum and copper, the NF metals described below are of particular technical relevance.

Magnesium

At a density of approximately 1.75 g/cm³, magnesium is significantly lighter than aluminum. The low melting temperature makes it particularly well-suited for castings. Typical applications in automotive engineering include car wheel rims, housing units and profiles.

Titanium

Hard-to-machine titanium alloys feature great corrosion resistance and a high level of thermal stability. Due to the favorable ratio of density and strength, these alloys lend themselves well to the aviation industry (e.g. air compressor turbine blades). Another application area is medical technology (implants, e.g. hip prosthetics) due to biocompatibility.

Sintered metals

Metal powder sintering

Sintered metals are generally manufactured using near-net-shape pressing of metal powders. This allows complicated shapes to be made from sintered metals at low costs, ready-to-install immediately or only requiring little finishing. After near-net-shape pressing the metal powder, a permanent connection between the grains is established by a running diffusion process at temperatures between 60 and 80 % of the melting temperature.

Metal injection molding

During MIM (metal injection molding), a process variant of sintering, a component is formed by injection molding metal powder and plastic compounds. After removing lubricants and binders with chemicals or heat, the molded articles retain their characteristic properties just like during metal powder sintering.

In addition to its chemical composition, the degree of porosity determines the properties and application of sintered metals to a large extent (see Table 10 and Table 11).

Table 9: Other alloys and cast copper alloys

Material group	Typical materials	Tensile strength [MPa]	Yield strength [MPa]	Fatigue strength under reversed bending stress (reference values) [MPa]	Properties	Application examples
Tin alloy	SnSb12Cu6Pb	–	60	28	Low hardness, good corrosion resistance	Sliding bearings
Die-cast zinc	ZP0410	330	250	80	Excellent casting characteristics, dimensional stability, surface quality	Thin-walled, low-stressed castings
Heat-conductor alloy	NiCr8020	650	–	–	High electrical resistance, high temperature stability	Heating filament
Resistance alloy	CuNi44	420	–	–	Low temperature coefficient, high oxidation stability	Resistors, potentiometer, heating filament
Cast tin bronze	CuSn10-C-GS	250	160	90	Good corrosion, resistance to wear	Fittings, pump housings
Red brass	CuSn7Zn4Pb7-C-GZ	260	150	80	Seawater resistant, good machinability, limp-home characteristics	Friction bearings, bushings

References for metallic materials

[1] DIN 30910-3: Sintered metal materials – Sintered-material specifications – Part 3: Materials for bearings and structural parts with bearing properties.

[2] DIN 30910-4: Sintered metal materials – Sintered-material specifications – Part 4: Materials for structural parts.

Table 10: Materials for bearings and structural parts with bearing properties¹

		Permissible ranges					Representative examples				
Material	Material code	Density ρ	Chemical composition in % by mass	Radial breaking strength	Hardness	Density ρ	Chemical composition in % by mass	Radial breaking strength	Compressive yield point	Hardness	Thermal conductivity
	Sint-	g/cm ³	Percent	$K^{(2)}$ N/mm ²	HB	g/cm ³	Percent	$K^{(2)}$ N/mm ²	$\delta_{0.02}$ N/mm ²	HB ⁽²⁾	λ W/mK
Sintered iron	A 00	5.6 to 6.0	<0.3 C; <1.0 Cu;	>150	>25	5.9	<0.2 others; rest Fe	160	130	30	37
	B 00	6.0 to 6.4	<2 others; rest Fe	>180	>30	6.3		190	160	40	43
	C 00	6.4 to 6.8		>220	>40	6.7		230	180	50	48
Sintered steel containing Cu	A 10	5.6 to 6.0	<0.3 C; 1 to 5 Cu;	>160	>35	5.9	2.0 Cu;	170	150	40	36
	B 10	6.0 to 6.4	<2 others; rest Fe	>190	>40	6.3	<0.2 others; rest Fe	200	170	50	37
	C 10	6.4 to 6.8		>230	>55	6.7		240	200	65	42
Sintered steel containing Cu and C	B 11	6.0 to 6.4	0.4 to 1.5 C; 1 to 5 Cu; <2 others; rest Fe	>270	>70	6.3	0.6 C; 2.0 Cu; <0.2 others; rest Fe	280	160	80	28
Sintered steel containing higher Cu	A 20	5.8 to 6.2	<0.3 C; 15 to 25 Cu;	>180	>30	6.0	20 Cu;	200	140	40	41
	B 20	6.2 to 6.6	<2 others; rest Fe	>200	>45	6.4	<0.2 others; rest Fe	220	160	50	47
Sintered steel containing higher Cu and C	A 22	5.5 to 6.0	0.5 to 3.0 C;	>120	>20	5.7	2.0 C ⁽³⁾ ; 20 Cu;	125	100	25	30
	B 22	6.0 to 6.5	15 to 25 Cu; <2 others; rest Fe	>140	>25	6.1	<0.2 others; rest Fe	145	120	30	37
Sintered bronze	A 50	6.4 to 6.8	<0.2 C; 9 to 11 Sn;	>120	>25	6.6	10 Sn;	140	100	30	27
	B 50	6.8 to 7.2	<2 others; rest Cu	>170	>30	7.0	<0.2 others; rest Cu	180	130	35	32
	C 50	7.2 to 7.7		>200	>35	7.4		210	160	45	37
Sintered bronze graphitic ⁽⁴⁾	A 51	6.0 to 6.5	0.5 to 3.0 C;	>100	>20	6.3	1.5 C ⁽⁴⁾ ; 10 Sn;	120	80	20	20
	B 51	6.5 to 7.0	9 to 11 Sn;	>150	>25	6.7	<0.2 others; rest Cu	155	100	30	26
	C 51	7.0 to 7.5	<2 others; rest Cu	>170	>30	7.1		175	120	35	32

¹ According to "Material Specification Sheets for Sintered Metals"; DIN 30910-3, 2004 edition [1]. ² Measured on calibrated bearings 10/16 Ø. 10.³ C is primarily present as free graphite. ⁴ C is present as free graphite.

Table 11 (continued): Sintered metals for structural parts¹

Permissible ranges				Representative examples							
Material	Material code	Density ρ	Chemical composition in % by mass	Hard-ness	Density ρ	Chemical composition in % by mass	Tensile strength R_m	Yield point $R_{p\ 0.1}$	Elonga-tion at break A	Hard-ness	E- module $E \cdot 10^3$ N/mm ²
Sintered steel containing Cu, Ni, Mo and C	Sint-		Percent	HB		Percent				HB	
	C 39 D 39	6.4 to 6.8 6.8 to 7.2	0.3 to 0.9 C; 1 to 3 Cu; 1 to 5 Ni; <0.6 Mo; <2 others; rest Fe	>90 >120	6.6 6.9	0.5 C; 1.5 Cu; 4.0 Ni; 0.5 Mo; <0.5 others; rest Fe	480 560	350 380	1 2	140 160	100 130
Stainless sintered steel AISI 316	C 40 D 40	6.4 to 6.8 6.8 to 7.2	<0.08 C; 10 to 14 Ni; 2 to 4 Mo; 16 to 19 Cr; <2 others; rest Fe	>95 >125	6.6 6.9	0.06 C; 13 Ni; 2.5 Mo; 18 Cr; <0.5 others; rest Fe	330 400	250 320	1 2	110 135	100 130
	AISI 430	C 42	6.4 to 6.8	>140	6.6	0.06 C; 18 Cr; <0.5 others; rest Fe	420	330	1	170	100
AISI 410	C 43	6.4 to 6.8	<0.3 C; 11 to 13 Cr; <2 others; rest Fe	>165	6.6	0.2 C; 13 Cr; <0.5 others; rest Fe	510	370	1	180	100
Sintered bronze	C 50 D 50	7.2 to 7.7 7.7 to 8.1	9 to 11 Sn; <2 others; rest Cu	>35 >45	7.4 7.9	10 Sn; <0.5 others; rest Cu	150 220	90 120	4 6	40 55	50 70
Sintered aluminum AlCuMgSi	E 73	2.55 to 2.65	4 to 6 Cu; <1 Mg; <1 Si; <2 other; rest Al.	>55	2.58 ² 2.58 ³	4.5 Cu; 0.6 Mg; 0.7 Si; <0.5 other; rest Al.	180 285	150 ⁴ n.c.	1 <0.5	65 90	55 ⁵ 55 ⁵
	F 75	2.60 to 2.66	2 to 3 Cu; <1 Mg; 13 to 16 Si; <2 other; rest Al. 1.5 to 2.0 Cu; 2.2 to 2.8 Mg; 5.6 to 6.4 Zn; <2 other; rest Al.	>70	2.63 ² 2.63 ³	2.5 Cu; 0.5 Mg; 14 Si; <0.5 other; rest Al.	200 300	180 ⁴ n.c.	<0.5 <0.5	90 125	78 ⁵ 78 ⁵
AlZnMgCu	F 77	2.74 to 2.78		>90	2.78 ² 2.78 ³	1.6 Cu; 2.6 Mg; 6.0 Zn; <0.5 other; rest Al.	300 450	190 ⁴ 230 ⁴	3 1.5	100 155	68 ⁵ 68 ⁵

¹ According to "Material Specification Sheets for Sintered Metals"; DIN 30910-4, 2010 edition [2].² T1 a sintered state and stored at room temperature for at least five days. ³ T6 solution annealed and artificially aged.⁴ Yield strength $R_{p0.2}$. ⁵ Determining the E-modulus using ultrasound.

n.c. not calculated.

EN metallurgy standards

Standardization of steels

(In accordance with DIN EN 10020, [1])
Steel is defined as an iron alloy, usually with a carbon content of $\leq 2\%$. Ferrous materials with a higher carbon content are usually classified as cast iron. Steels are classified in three classes, "unalloyed steels", "stainless steels" and "other alloyed steels".

Unalloyed steels

Unalloyed steels that do not reach the defined minimum alloying element content are then divided into unalloyed high-grade steels and unalloyed stainless steels. For unalloyed high-grade steels, defined requirements apply, such as those regarding toughness and malleability.

Unalloyed stainless steels feature a higher degree of purity and, therefore, improved properties such as high yield strength, hardenability, good toughness and weldability. They are usually intended for quenching and drawing or for surface hardening.

Stainless steels

Stainless steels feature a chrome content of at least 10.5 percent by mass and a carbon content of less than 1.2 percent by mass. They are subdivided further by nickel content (more or less than 2.5 percent by mass) and by the main properties, including corrosion resistance, heat resistance and temperature stability.

Other alloyed steels

These include steels with requirements in regards to aspects such as toughness, particle size or malleability. They are usually not intended for quenching and drawing or surface hardening. A distinction is made between alloyed high-grade steels and alloyed stainless steels.

Designation system for steels with material abbreviations

(In accordance with DIN EN 10027-1, [2])
The material abbreviations for steels are divided into two groups.

- Group 1: Material abbreviations that feature information on use and mechanical and physical properties.
- Group 2: Material abbreviations that contain information on chemical composition.

Abbreviated names of Group 1

These material abbreviations include notes on use and mechanical or physical properties. The prefix G indicates that it involves a cast steel material:

S, GS	For general steel construction,
P, GP	For pressure-vessel construction,
L	For pipeline construction,
E	Engineering steels,
B	Concrete reinforcing steels,
Y	Prestressing steels,
R	Rail steels,
H	Flat products made of higher-strength steels for cold forming,
D	Flat products for cold forming,
T	Packaging steel sheet and strip,
M	Electrical steel sheet and strip.

Examples:

S235JR	
S	Generally for steel construction,
235	Yield strength in MPa,
JR	Notch impact work 27 J at 20 °C.

HC240LA

H	High-strength steel,
C	Cold-rolling,
240	Minimum yield strength in MPa,
LA	Low-alloyed.

Abbreviated names of Group 2

These material abbreviations contain references to the chemical composition. The prefix G indicates that this is a cast steel material:

C, GC	Unalloyed steels (Mn < 1%),
G	Unalloyed steels (Mn low-alloyed steels),
X, GX	High-alloyed steels,
PM	Powder metallurgy,
HS	High-speed steel.

Examples:**C85S**

C Unalloyed steel ($< 1\%$ Mn),
85 0.01 times the C content,
 i.e. 0.85% C,
S For springs.

42CrMo4

42 Low-alloyed steel ($Mn \geq 1\%$)
 0.01 times the C content,
 i.e. 0.42% C,
Cr 4 times the Cr content
 in percent, i.e. 1% Cr,
Mo Non-specific percentage of
 alloyed constituents, Mo ($< 1\%$).

X5CrNi18-10

X High-alloyed steel,
5 0.05% carbon,
18 18% Cr,
10 10% Ni.

HS 7-4-2-5

HS High-speed steel
 Percentage of alloyed constituents in whole percentages in the order of tungsten – molybdenum – vanadium – cobalt:
7 7% tungsten,
4 4% molybdenum,
2 2% vanadium,
5 5% cobalt.

In addition, special requirements can be defined, such as for the type of coating or notes on the treatment condition. Example information could read as follows:

+H With hardenability,
+CU With copper coating,
+Z Hot-dip galvanized,
+ZE Electrolytically galvanized,
+C Work-hardened,
+M Thermomechanically shaped,
+Q Quenched,
+U Untreated.

Designation system for steels based on the numbering system

(In accordance with DIN EN 10027-2, [3])
 In addition to their material abbreviation, all steels are defined by a material number in accordance with the following structure:

Major group + steel group + counter.

Material major group number:

0 Pig iron, ferro-alloys,
1 Steel,
2 Nonferrous heavy metals,
3 Light metals,
4 to 8 Non-metallic materials,
9 Unassigned for internal use.

Steel group number (selection):

00 Unalloyed ordinary low-carbon steels,
01 to 07 Unalloyed high-grade steels,
10 to 18 Unalloyed stainless steels,
40 to 49 Chemically stable alloying steels,
20 to 29 Alloying tool steels.

Example:

1.4301
 (material abbreviation: X5CrNi18-10)
1 Steel,
43 Stainless steel,
 with more than 2.5% Ni,
 without any Mo, Nb or Ti,
01 Count number.

Standardization of cast-iron materials

The structure and form of the carbon (carbide or graphite) and the graphite structure are important to the properties. In standardization, a distinction is made between the following four groups of cast iron:

- Lamellar graphite cast iron (DIN EN 1561, [4]),
- Malleable cast iron (DIN EN 1562, [5]),
- Spheroidal graphite cast iron (DIN EN 1563, [6]),
- Ausferritic spheroidal graphite cast iron (DIN EN 1564, [7]).

There are two systems for designating types of cast iron: In accordance with material numbers or abbreviations.

Designation system for cast iron with abbreviations

(In accordance with DIN EN 1560, [8])

The designation is alphanumeric and consists of up to six individual positions. The first position is EN (European standard). This is followed by G (cast) and J (iron). The character for denoting the graphite structure is in the third position:

L	Lamellar,
S	Spheroidal,
M	Temper carbon,
V	Vermicular,
N	Graphite-free,
	ledeburitic cast chilled iron,
Y	Special structure.

Notes on the microstructure and macrostructure can be placed at position 4 if necessary, such as (as an excerpt):

A	Austenite,
F	Ferrite,
Q	Quenched.

The fifth position contains information on tensile strength, impact energy with test temperature or for hardness. Alternatively, the fifth position can specify the chemical composition.

Finally, the sixth position contains additional requirements such as:

D	As-cast casting,
H	Heat-treated casting.

Examples:

EN-GJL-150

EN	European standard,
GJ	Cast iron,
L	Lamellar graphite,
150	Tensile strength in MPa.

EN-GJV-HV400

EN	European standard,
GJ	Cast iron,
V	Vermicular graphite,
HV400	Vickers hardness.

EN-GJS-SiMo30-8

EN	European standard,
GJ	Cast iron,
S	Spheroidal graphite,
Si	Silicon content 3%,
Mo	Molybdenum content 0.8%.

Designation system for cast iron with material numbers

(In accordance with DIN EN 1560, [8])

The designation consists of a total of six positions. The first and second positions contain "5." followed by information on the graphite structure at the third position:

1	Lamellar,
2	Vermicular,
3	Spheroidal,
4	Temper carbon.

The fourth position describes the matrix structure (selection):

1	Ferrite,
3	Pearlite,
5	Austenite.

The fourth position is followed by a double-digit count number (00-99) that identifies the individual material.

Nonferrous-metal alloys

As for ferrous materials, the EN standard also features two options for describing and identifying each nonferrous metals (NF) material and its alloys: The first designation system uses chemical symbols (abbreviations) and the second uses a numerical designation system.

Unlike steels, NF metals are each identified with their own numerical designation system. Thus, the different properties and requirements of, for example, aluminum, copper or zinc alloys are better taken into account.

Designation system for NF metals with abbreviation

The EN standard denotes nonferrous metals (NF metals) according to the following basic system:

Example:

EN AW-Al Si1MgMn T6

EN	European standard,
----	--------------------

Code letter for metal:

A	Aluminum,
C	Copper,
M	Magnesium.

Code letter for processing:

W Wrought (wrought alloy),
C Casting (casting alloy).

Designation system with chemical symbols, e.g. AlSi1MgMn in this case:

Material condition, here, e.g.: T6.

The system lists the alloy metals in order of base metal with falling percentages.

Special numerical designation system for aluminum (Al) and aluminum alloys

For wrought Al products (Al and wrought Al alloys in accordance with DIN EN 573-1, [9]), the alloy is determined by four digits (Example 1); for Al casting alloys (in accordance with DIN EN 1706, [10]), by five digits (Example 2).

Example 1:

EN AW-6082 T6

EN European standard,
AW Wrought aluminum alloy,
6 Indicator 6 for alloy group (Al-Mg-Si alloys),
0 Original alloy (1, 2 for modifications),
82 Designation for the alloy with about 1% Si, 0.7% Mn and 0.9% Mg,
T6 Material condition T6 (solution-annealed and artificially aged).

Example 2:

EN AC -45200

EN European standard,
AC Casting aluminum alloy,
45 AlSi5Cu alloy group,
200 Number for individual alloy (here AlSi5Cu3Mn).

Code numbers for the alloy groups:

1 Pure aluminum,
2 With copper,
3 With manganese,
4 With silicon,
5 With magnesium,
6 With magnesium-silicon,
7 With zinc,
8 Other.

Material conditions (selection):

O Soft-annealed,
H Work-hardened,
H14 Work-hardened, 1/2 hard (for sheet metal),
T6 Solution-annealed and artificially aged.

References for EN metallurgy standards

- [1] DIN EN 10020: Definition and classification of grades of steel.
- [2] DIN EN 10027-1: Designation systems for steels – Part 1: Steel names.
- [3] DIN EN 10027-2: Designation systems for steels – Part 2: Numerical system.
- [4] DIN EN 1561: Founding – Grey cast irons.
- [5] DIN EN 1562: Founding – Malleable cast irons.
- [6] DIN EN 1563: Founding – Spheroidal graphite cast irons.
- [7] DIN EN 1564: Founding – Ausferritic spheroidal graphite cast irons.
- [8] DIN EN 1560: Founding – Designation system for cast iron – Material symbols and material numbers.
- [9] DIN EN 573-1: Aluminium and aluminium alloys – Chemical composition and form of wrought products – Part 1: Numerical designation system.
- [10] DIN EN 1706: Aluminium and aluminium alloys – Castings – Chemical composition and mechanical properties.

Magnetic materials

Materials which have ferromagnetic or ferrimagnetic properties are called magnetic materials and belong to one of two groups: metals (metals produced through melting or sintering) or nonmetallic inorganic materials. Composite materials such as soft magnetic composite materials and plastic-bonded permanent magnets are also playing an increasingly important role. They are characterized by their ability to exhibit a permanent external field or by their good magnetic flux conductivity (soft magnets).

In addition to ferromagnets and ferrimagnets, diamagnetic, paramagnetic, and antiferromagnetic materials also exist. They differ from each other in terms of their permeability μ , or the temperature dependence of their susceptibility κ . This variable gives the ratio of the magnetization of a substance to the magnetic field strength or excitation.

$$\mu_r = 1 + \kappa.$$

Classification

Ferromagnets and ferrimagnets

Both substances exhibit spontaneous magnetization which disappears at the Curie point (Curie temperature T_C). At temperatures above the Curie temperature, they behave like paramagnets. For susceptibility κ at $T > T_C$, the Curie-Weiss law applies as follows:

$$\kappa = \frac{C}{T - T_C},$$

where

C Curie constant,
 T Temperature in K.

The saturation induction of ferromagnets is higher than for ferrimagnets because all magnetic moments are aligned in parallel. In the case of ferrimagnets, on the other hand, the magnetic moments of the two sublattices are aligned antiparallel to one another. Nevertheless, these materials are magnetic because the magnetic moments of the two sublattices have different magnitudes.

Diamagnets

For diamagnets, susceptibility κ_{Dia} is independent of the temperature.

Paramagnets

For paramagnets, susceptibility κ_{Para} drops as temperature increases. In this case, the Curie law states:

$$\kappa_{\text{Para}} = \frac{C}{T}.$$

Antiferromagnets

As in the case of ferrimagnets, adjacent magnetic moments are aligned antiparallel with respect to one another. As they are of equal magnitude, the effective magnetization of the material is zero.

At temperatures above the Néel point (Néel temperature T_N), antiferromagnets behave like paramagnets. At $T > T_N$, the susceptibility is:

$$\kappa = \frac{C}{T + \Theta},$$

where

Θ Asymptotic Curie temperature.

Examples of antiferromagnets: MnO, MnS, FeCl₂, FeO, NiO, Cr, V₂O₃, V₂O₄.

Soft magnetic materials

The figures specified in Table 1 are from the applicable DIN standards (soft-magnetic metallic materials, DIN IEC 60404-8-6, [1]). Many material qualities defined in this standard relate to the materials in DIN 17405 (DC relays, [2]) and DIN-IEC 60740-2 (transformers and reactors, [3]).

Designation (composition)

"Code letter" "Number 1" "Number 2" – "Number 3".

The code letter specifies the alloy class: "A" pure iron, "C" silicon-iron (SiFe), "E" nickel-iron (NiFe), "F" cobalt-iron (CoFe).

Number 1 indicates the concentration of the main alloy element.

Number 2 defines the different curves: 1 indicates a round hysteresis loop, 2 indicates a rectangular hysteresis loop.

The significance of Number 3 following the hyphen varies according to the individual alloy. It indicates the minimum initial permeability $\mu_a/1,000$ in nickel alloys; with other alloys, it designates the maximum coercive field strength in A/m. The properties of these materials are strongly geometry-dependent and highly application-specific. The material data quoted in extracts from the standard can therefore provide only a very general overview of the properties of these materials.

Electrical steel sheet and electrical steel strip

Applications and properties

Packages used for transformers or as stator or rotor packages for electric motors are composed of electrical steel strips, typically by packaging individual sheet metal plates. Electrical steel strips are often provided as long strips (wide strips or split strips) in a wound shape (called coils), but they are also available as panels (electrical steel sheet).

Electrical steel strips typically consist of iron-silicon alloys ([4] and [5]); other materials such as aluminum and manganese may function as alloy constituents in small quantities. Densities between 7.65 and 7.85 g/cm³ result depending on the composition. They are characterized by low remagnetization loss, high polarization and high permeability compared to other soft magnetic ferrous materials. The static coercive field strength H_c is typically 100 to 300 A/m for A and K types (see Designations), the maximum permeability $\mu_{\max}$ is at about 5,000. For S and P types, the values are in the order of magnitude of 1 A/m for H_c and about 30,000 for $\mu_{\max}$.

Grain oriented electrical steel strip

Both grain oriented and non-oriented electrical steel strips exist. Grain oriented electrical steel strips feature a crystal structure (cubic structure), meaning the grains are oriented in a preferred direction by rolling and annealing steps during manufacturing. This gives the strip material a preferred magnetic direction in the strip direction. The magnetic properties depend on the direction (anisotropic magnetic properties). Grain oriented electrical

steel strip is primarily used in components where exceptionally good magnetic flux in a specific direction comes into play (e.g. transformers, choke coils and converters).

Non-oriented electrical steel strip

In non-oriented electrical steel strip, however, there is not a structure. In other words, the crystallographic orientation of the grains is almost randomly arranged. The magnetic properties do not depend on the direction (isotropic magnetic properties). For this reason, non-oriented electrical steel strip is used in components where magnetic flux is not limited to a preferred direction (e.g. electric motors and alternators) [6].

Non-oriented electrical steel strips are usually delivered in a finally annealed (fully-finished) state. In other words, final annealing for optimizing the magnetic properties takes place at the manufacturer of the electrical steel strips. For electrical steel strips that are not finally annealed (semi-finished), final annealing still has to be carried out. This is typically done to the finished laminated core.

Coating

Electrical isolation of the individual layers suppresses eddy currents for alternating-field magnetization, reducing remagnetization losses.

Identifications

The designations for electrical steel sheets and electrical steel strips (ESS) is defined in DIN EN 10027-1 [7]:

Code letter 1 Number 1 – Number 2 Code letter 2 (see below for an example).

The first code letter is "M" for all varieties. Number 1 is one hundredfold the maximum magnetic reversal loss at 50 Hz and 1.5 T (types A and K) or 1.7 T (types S and P) in W/kg. Number 2 is one hundredfold the product's nominal depth in mm.

Table 2: Properties of electrical steel strips and electrical sheet steel

Sheet grade		Nominal thickness mm	Maximum remagnetization loss in W/kg at 50 Hz and activation of			Minimum magnetic polarization in tesla (T) in an AC field at field strength H in A/m			Applications
Abbreviated name	Material number		1.0 T	1.5 T	1.7 T	2,500	5,000	10,000	
M270-35A	1.0801	0.35	1.10	2.7	—	1.49	1.60	1.70	Electric motors
M330-35A	1.0804	0.35	1.30	3.3	—	1.49	1.60	1.70	
M330-50A	1.0809	0.50	1.35	3.3	—	1.49	1.60	1.70	
M530-50A	1.0813	0.50	2.30	5.3	—	1.56	1.65	1.75	
M800-50A	1.0816	0.50	3.60	8.0	—	1.60	1.70	1.78	
M400-65A	1.0821	0.65	1.70	4.0	—	1.52	1.62	1.72	
M1000-65A	1.0829	0.65	4.40	10.0	—	1.61	1.71	1.80	
M800-100A	1.0895	1.00	3.60	8.0	—	1.56	1.66	1.75	
M1300-100A	1.0897	1.00	5.80	13.0	—	1.60	1.70	1.78	
M340-50K	1.0841	0.50	1.42	3.4	—	1.54	1.62	1.72	Low-power motors for industrial and household appliances (e.g. washing machine motors, microwave transformers, refrigerator compressors)
M560-50K	1.0844	0.50	2.42	5.6	—	1.58	1.66	1.76	
M660-50K	1.0361	0.50	2.80	6.6	—	1.62	1.70	1.79	
M1050-50K	1.0363	0.50	4.30	10.5	—	1.57	1.65	1.77	
M390-65K	1.0846	0.65	1.62	3.9	—	1.54	1.62	1.72	
M630-65K	1.0849	0.65	2.72	6.3	—	1.58	1.66	1.76	
M800-65K	1.0364	0.65	3.30	8.0	—	1.62	1.70	1.79	
M1200-65K	1.0366	0.65	5.00	12.0	—	1.57	1.65	1.77	
						At field strength $H = 800$ A/m			Converters, transformers, choke coils
M140-30S	1.0862	0.30	—	0.92	1.40	1.78			
M150-30S	1.0861	0.30	—	0.97	1.50	1.75			
M111-30P	1.0881	0.30	—	—	1.11	1.88			

Code letter 2 provides type data:

- “A” cold rolled non-oriented electrical steel sheet and strip delivered in the fully processed state (DIN EN 10106, [8]).
- “S” conventional grain-oriented electrical steel sheet or “P” grain-oriented electrical steel strip with high permeability, both types are delivered in a fully processed state (DIN EN 10107, [9]).
- “K” cold rolled electrical non-alloy and alloy steel sheet and strip delivered in the semi-processed state (DIN EN 10341, [10]).

Example: An electrical steel strip designated M330-35A is a non-oriented electrical steel strip in a finally annealed state, exhibits a remagnetization loss of 3.3 W/kg at 50 Hz and 1.5 T and has a thickness of 0.35 mm.

Materials for transformers and reactors (DIN IEC 740-2)

These materials comprise the alloy classes C21, C22, E11, E31 and E41 from the standard for soft-magnetic materials (DIN IEC 60404-8-6). The standard essentially contains the minimum values for core-sheet permeability for specified core-sheet sections (YEI, YED, YEE, YEL, YUI, and YM).

Materials for direct-current relays (DIN 17405)

Designation comprises a letter and number combination:

- a) Code letter "R" (relay material).
- b) Code letters for identifying alloying constituents:
Fe = unalloyed, Si = silicon steels,
Ni = nickel steels or alloys.
- c) Code number for maximum coercive field strength.
- d) Code letter for the stipulated delivery state: "U" = untreated, "GB" = malleable pre-annealed, "GT" = pre-annealed for deep-drawing, "GF" = final-annealed.

DIN IEC 60404-8-10 essentially contains the limit deviations for magnetic relay materials based on iron and steel. The designation code defined in this standard is as follows (example: M 80 TH):

- Code letter "M".
- Permitted maximum value for coercive field strength in A/m.
- Code letter for material composition: "F" = pure iron, "T" = steel alloy, "U" = unalloyed steel.
- Code letter for delivery state: "H" = hot-rolled, "C" = cold-rolled or cold-drawn.

Sintered metals for soft magnetic components (DIN IEC 60404-8-9)

Designation comprises a letter and number combination:

- Code letter "S" for sintered materials.
- Hyphen followed by the identifying alloy elements, i.e. Fe plus P, Si, Ni, or Co if necessary.
- The maximum permitted coercive field strength in A/m follows the second hyphen.

Soft magnetic ferrite cores (previously DIN 41280, [11])

Soft-magnetic ferrites are formed parts made of a sintered material with the general formula $MO \cdot Fe_2O_3$, where M is one or more of the bivalent metals Cd, Co, Ca, Mg, Mn, Ni, Zn.

Designation comprises a letter and number combination:

The various types of soft-magnetic ferrites are classified in groups according to nominal initial permeability and are designated by capital letters. Additional numbers may be used to further subdivide them into subgroups; these numbers have no bearing on material quality.

The coercive field strength H_c of soft ferrites is usually in the range of 4 to 500 A/m. Based on a field strength of 3,000 A/m, induction B is in the range of 350 to 470 mT.

Powder composite materials

Powder composite materials are not yet standardized, but are becoming increasingly more important. They consist of ferromagnetic metal powder (iron or an alloy) and an organic or inorganic grain-boundary phase as a "binder". They are manufactured in much the same way as sintered metals. The individual manufacturing stages are:

- Mixing the starting materials (metal powder and binder).
- Shaping by injection-molding, extruding or pressing.
- Heat treatment below the sinter temperature ($< 600^\circ\text{C}$).

Depending on the shaping process, binder type and the amount of binder used, it is possible to optimize the material to achieve high saturation polarization, higher permeability or high resistivity.

They are used primarily in fields in which all the above-mentioned characteristics are important, and where no excessively high demands are placed on mechanical strength and machinability. These fields currently consist of quick-acting actuators for diesel-injection engineering and high-speed small electric motors for motor vehicles.

Table 3: Materials for transformers and reactors

Core-sheet permeability for alloy classes C21, C22, E11, E31, and E41 for core-sheet section YE11.

Minimum core-sheet permeability μ_{lam} (min)									
IEC designation	C21-09 Thickness in mm		C22-13 Thickness in mm		E11-60 Thickness in mm				
	0.3 to 0.38	0.15 to 0.2	0.3 to 0.38	0.15 to 0.2	0.3 to 0.38	0.15 to 0.2	0.1	0.05	
YE1 1	-10	630	630	1,000	14,000	18,000	20,000	20,000	20,000
	13	800	630	1,000	18,000	20,000	22,400	20,400	20,400
	14	800	630	1,000	18,000	22,400	22,400	22,400	22,400
	16	800	630	1,000	20,000	22,400	25,000	22,400	22,400
	18	800	630	1,000	22,400	25,000	25,000	25,000	22,400
	20	800	630	1,120	22,400	25,000	25,000	25,000	25,000
	22	800	630	1,120	22,400	25,000	25,000	25,000	25,000
IEC designation	E11-100 Thickness in mm		E31-04 Thickness in mm		E31-06 Thickness in mm				
	0.3 to 0.38	0.15 to 0.2	0.3 to 0.38	0.15 to 0.2	0.3 to 0.38	0.15 to 0.2	0.1	0.05	
	-10	18,000	25,000	2,800	3,150	3,150	4,000	4,500	5,000
	13	20,000	28,000	2,800	3,150	3,150	4,000	4,500	5,000
	14	22,400	28,000	2,800	3,150	3,150	4,000	4,500	5,000
	16	25,000	31,500	2,800	3,150	3,150	4,500	5,000	5,000
	18	25,000	31,500	3,150	3,150	3,550	4,500	5,000	5,000
IEC designation	E31-10 Thickness in mm		E41-02 Thickness in mm		E41-03 Thickness in mm				
	0.3 to 0.38	0.15 to 0.2	0.3 to 0.38	0.15 to 0.2	0.3 to 0.38	0.15 to 0.2	0.1	0.05	
	-10	5,600	6,300	1,600	1,800	1,800	2,000	2,240	2,240
	13	6,300	7,100	1,800	1,800	2,000	2,000	2,240	2,240
	14	6,300	7,100	1,800	1,800	2,000	2,000	2,240	2,240
	16	6,300	7,100	1,800	1,800	2,000	2,000	2,240	2,240
	18	7,100	7,100	1,800	1,800	2,000	2,000	2,240	2,240
	20	7,100	7,100	1,800	2,000	2,000	2,000	2,500	2,240

Table 4: Materials for DC relays

Material type		Alloying constituents	Dens-ity ¹⁾ ρ	Hard-ness ¹⁾	Rema-nence ¹⁾ T (tesla)	Permeability ¹⁾ μ_{max}	Electr. resist-ivity ¹⁾ $\Omega \cdot \text{mm}^2$ m	Coercive field strength A/m max.	Magnetic polarization in T (Tesla) min. at field strength H in A/m						Proper-ties, applica-tion examples	
Abbreviated Name	Material number	by mass %	g/cm ³	HV					20	50	100	200	300	500	1,000	4,000
Unalloyed steels																
RFe 160	1.1011	-	7.85	max. 150	-	-	0.15	160	-	-	-	-	1.15	1.30	-	1.60
RFe 80	1.1014				1.10	-	0.15	80	-	1.10	1.20	1.30	1.45	1.60		
RFe 60	1.1015				1.20	-	0.12	60	-	1.15	1.25	1.35	1.45	1.60		
RFe 20	1.1017				1.20	} $\approx 20,000$	0.10	20	-	1.15	1.25	1.30	1.40	1.45	1.60	
RFe 12	1.1018				1.20		0.10	12	-	1.15	1.25	1.30	1.40	1.45	1.60	
Silicon steels																
RSi 48	1.3840	2.5	7.55	130	0.50	-	0.42	48	-	-	0.60	-	1.10	1.20	-	1.50
RSi 24	1.3843	-	-	-	1.00	$\approx 20,000$	-	24	-	-	1.20	-	1.30	1.35	-	1.50
RSi 12	1.3845	4 Si	7.75	200	1.00	$\approx 10,000$	0.60	12	-	-	1.20	-	1.30	1.35	-	1.50
Nickel steels and nickel alloys																
RNi 24	1.3911	≈ 36 Ni	8.2	130 to 180	0.45	$\approx 5,000$	0.75	24	0.20	0.45	0.70	-	0.90	1.0	-	1.18
RNi 12	1.3926	≈ 50 Ni	8.3	130 to 180	0.60	$\approx 30,000$	0.45	12	0.50	0.90	1.10	-	1.25	1.35	-	1.45
RNi 8	1.3927	≈ 50 Ni	8.3	130 to 180	0.60	30,000 to 100,000	0.45	8	0.50	0.90	1.10	-	1.25	1.35	-	1.45
RNi 5	2.4596	70 to 80 Ni	8.7	120 to 170	0.30	$\approx 40,000$	0.55	5	0.50	0.65	0.70	-	-	-	-	0.75
RNi 2	2.4595	small quantities Cu, Cr, Mo	8.7	120 to 170	0.30	$\approx 100,000$	0.55	2	0.50	0.65	0.70	-	-	-	-	0.75

¹⁾ Standard values.

Table 5: Sintered metals for soft-magnetic components

Material	Characteristic alloy elements (except Fe) % by mass	Sinter Density ρ_s	Porosity p_s	Maximum coercive field strength $H_{c(max)}$	Magnetic polarization in tesla (T) at field strength H in A/m			Maximum permeability	Vickers hardness	Volume Electrical resistivity ρ
Abbreviated name	%	g/cm ³	%	A/m	500	5,000	15,000	80,000	HV5	$\mu\Omega m$
S-Fe-175	—	6.6	16	175	0.70	1.10	1.40	1.55	50	0.15
S-Fe-170	—	7.0	11	170	0.90	1.25	1.45	1.65	60	0.13
S-Fe-165	—	7.2	9	165	1.10	1.40	1.55	1.75	70	0.12
S-FeP-150	≈ 0.45 P	7.0	10	150	1.05	1.30	1.50	1.65	95	0.20
S-FeP-130	≈ 0.45 P	7.2	8	130	1.20	1.45	1.60	1.75	105	0.19
S-FeSi-80	≈ 3 Si	7.3	4	80	1.35	1.55	1.70	1.85	170	0.45
S-FeSi-50	≈ 3 Si	7.5	2	50	1.40	1.65	1.70	1.95	180	0.45
S-FeNi-20	≈ 50 Ni	7.7	7	20	1.10	1.25	1.30	1.30	70	0.50
S-FeNi-15	≈ 50 Ni	8.0	4	15	1.30	1.50	1.55	1.55	85	0.45
S-FeCo-100	≈ 50 Co	7.8	3	100	1.50	2.00	2.10	2.15	190	0.10
S-FeCo-200	≈ 50 Co	7.8	3	200	1.55	2.05	2.15	2.20	240	0.35

Table 6: Soft magnetic ferrites

Ferrite type	Initial permeability ¹⁾ μ_i $\pm 25\%$	Specific loss factor $\tan \delta(\mu_i^{(2)})$		Specific Power loss ³⁾ mW/g	Amplitude permeability ⁴⁾ μ_a	Curie temperature ⁵⁾ θ_c °C	Frequency for $0.8 \cdot \mu_i^{(6)}$ MHz	Properties, application examples
Materials in largely open magnetic circuits								
C 1/12	12	350	100	—	—	> 500	400	Initial permeability. Compared to metallic magnetic materials high specific resistance (100 to 10 ⁵ Ω · m, metals 10 ⁻⁷ to 10 ⁻⁶ Ω · m), therefore low eddy-current losses. Telecommunications (coils, transformers).
D 1/50	50	120	10	—	—	> 400	90	
F 1/250	250	100	3	—	—	> 250	22	
G 2/600	600	40	1	—	—	> 170	6	
H 1/1,200	1,200	20	0.3	—	—	> 150	2	
Materials in largely closed magnetic circuits								
E 2	60 to 160	80	10	—	—	> 400	50	
G 3	400 to 1,200	25	1	—	—	> 180	6	
J 4	1,600 to 2,500	5	0.1	—	—	> 150	1.5	
M 1	3,000 to 5,000	5	0.03	—	—	> 125	0.4	
P 1	5,000 to 7,000	3	0.01	—	—	> 125	0.3	
Materials for power applications								
W 1	1,000 to 3,000	—	—	45	1,200	> 180	—	
W 2	1,000 to 3,000	—	—	25	1,500	> 180	—	

¹⁾ Nominal values. ²⁾ $\tan \delta/\mu_i$ denotes the frequency-dependent material losses at a low flux density ($B < 0.1$ mT).

³⁾ Losses at high flux density. Measured preferably at: $f = 25$ kHz, $B = 200$ mT, $\theta = 100^\circ \text{C}$.

⁴⁾ Permeability when subjected to a strong sinusoidal magnetic field. Measured at: $f \leq 25$ kHz, $B = 320$ mT, $\theta = 100^\circ \text{C}$.

⁵⁾ Curie temperature θ_c in this table is the temperature at which initial permeability μ_i drops below 10 % of its value at 25°C .

⁶⁾ Standard values.

Permanent-magnet materials

(DIN 17410, replaced by DIN IEC 60404-8-1)
 If chemical symbols are used in the abbreviated names of the materials, they refer to the primary alloying constituents of the materials. The numbers in the material abbreviation before the slash denote the $(BH)_{\max}$ value in kJ/m^3 and those after the slash denote one tenth of the H_cJ value in kA/m (rounded values). Permanent magnets with binders are indicated by a final p.

Designation by material abbreviation or code number

Structure of code numbers

(DIN IEC 60404-8-1:2005-8):

Code letter (group)

+ 1st digit (material type),

+ 2nd digit of 0 (isotropic)

or 1 (anisotropic),

+ 3rd digit (various quality levels).

R – Hard magnetic alloys,

such as R1: Aluminum-nickel-cobalt-iron-titanium alloys (AlNiCo).

S – Hard magnetic ceramic materials, such as S1: Hard magnetic ferrite.

U – Combined hard magnetic materials, such as

– U1: Combined aluminum-nickel-cobalt-iron-titanium magnets (AlNiCo).

– U2: Combined rare earth-cobalt magnets (RECo).

– U3: Combined neodymium-iron-boron magnets.

– U4: Combined hard ferrites.

Comparison of permanent and soft magnets

Figure 1 shows the magnetic characteristic value ranges of some crystalline materials in widespread use. The values of soft magnetic materials are compared to the values of hard magnets (permanent magnets).

Table 7: Permanent magnet materials

Bosch grades BTMT (not standardized)

Material	Density ¹⁾ ρ	$(BH)_{\max}$ ²⁾	Remanence ²⁾ B_r	Coercive field strength ²⁾	
Abbreviated name	g/cm ³	kJ/m ³	mT	of flux density H_{CB} kA/m	of polarization H_{CJ} kA/m
RBX HC 370	4.7 to 4.9	25	360	270	390
RBX HC 380		28	380	280	370
RBX 380 K		28	380	280	300
RBX 400		30	400	255	260
RBX 400 K		31	400	290	300
RBX HC 400		29	380	285	355
RBX 420		34	420	255	270
RBX 410 K		33	410	305	330
RBX HC 410		30	395	290	340
RBX 420 S		35	425	260	270
RBX HC 400 N		28	380	280	390

¹⁾ Standard values. ²⁾ Minimum values.

Table 8: Permanent magnet materials

Material	Material number		Chemical composition ¹⁾ % by weight							Density ρ ¹⁾	$(BH)^{max}_{30}$	Remanence B_r ²⁾	Coercive field strength ²⁾ of flux density H_{CB}		Rel. permanent permeability ¹⁾	Curie temp. ¹⁾ T_c	Temp. coeff. of polar. $TK (J_s)_{1/3}$	Temp. coeff. of coerciv. $TK (H_c)_{1/3}$	Manufacture, processing, applications
Abbreviated name	DIN	IEC	Al	Co	Cu	Nb	Ni	Ti	Fe	ρ g/cm ³	$(BH)^{max}_{30}$ kJ/m ³	mT	H_{CB} kA/m	H_{CJ} kA/m	μ_p	K	%K	%K	
Metallic magnets																			
Isotropic																			
AlNiCo 9/5	1.3728	R 1-0-3	11 to 13	0 to 5	2 to 4	—	21 to 28	0 to 1	Rest	6.8	9.0	550	44	47	4.0 to 5.0	1,030	—	+ 0.03 to - 0.07	For magnets with binders: pressing or injection molding. Processing: grinding. Application: max. 400 to 500 °C.
AlNiCo 18/9	1.3756	—	6 to 8	24 to 34	3 to 6	—	13 to 19	5 to 9	Rest	7.2	18.0	600	80	86	3.0 to 4.0	to 1,180	- 0.02		
AlNiCo 7/8p	1.3715	R 1-2-3	6 to 8	24 to 34	3 to 6	—	13 to 19	5 to 9	Rest	5.5	7.0	340	72	84	2.0 to 3.0	1,180			
AlNiCo 35/5	1.3761	—	8 to 9	23 to 26	3 to 4	0 to 1	13 to 16	—	—	7.2	35.0	1,120	47	48	3.0 to 4.5	1,030			
AlNiCo 44/5	1.3757	R 1-1-2	8 to 9	23 to 26	3 to 4	0 to 1	13 to 16	—	—	7.2	44.0	1,200	52	53	2.5 to 4.0	to 1,030	+ 0.03 to - 0.07		
AlNiCo 52/6	1.3759	—	8 to 9	23 to 26	3 to 4	0 to 1	13 to 16	—	Rest	7.2	52.0	1,250	55	56	1.5 to 3.0	to 1,180	- 0.02		
AlNiCo 60/11	1.3763	R 1-1-6	6 to 8	35 to 39	2 to 4	0 to 1	13 to 15	4 to 6	—	7.2	60.0	900	110	112	1.5 to 2.5	1,180			
AlNiCo 30/14	1.3765	—	6 to 8	38 to 42	2 to 4	0 to 1	13 to 15	7 to 9	—	7.2	30.0	680	136	144	1.5 to 2.5	1,180			
PLiCo 60/40	2.5210	R2-0-1	Pt	Co						15.5	60	600	350	400	1.1	800	- 0.01 to - 0.02		
FeCoVCr 11/2	2.4570	R 3-1-3	V	Co	Cr	Fe				—	11.0	800	24	24	2.0 to 8.0	1,000	- 0.01		
FeCoVCr 4/1	2.4571	—	8 to 15	51 to 54	0 to 4	Rest				—	4.0	1,000	5	5	9.0 to 25.0	1,000			
RECo – Magnets of type RECo ₅																			
RECo 80/80	—	R 5-1-1	Typically MMCos (MM = Cer composition metal)							8.1	80	650	500	800	1.05	1,000	- 0.05	- 0.3	
RECo 120/96	—	R 5-1-2	Typically SmCos, typically (SmPr) Co ₅							8.1	120	770	590	960	1.05	1,000	- 0.05	- 0.3	
RECo 160/80	—	R 5-1-3								8.1	160	900	640	800	1.05	1,000	- 0.05	- 0.3	
RECo – Magnets of type RE ₂ Co ₁₇																			
RECo 165/50	—	R 5-1-11	Typically MMCos (MM = Cer composition metal)							8.2	165	950	440	500	1.1	1,100	0.03	- 0.02	
RECo 180/90	—	R 5-1-13	Typically SmCos, typically (SmPr) Co ₅							8.2	180	1,000	680	900	1.1	1,100	0.03	- 0.02	
RECo 190/70	—	R 5-1-14								8.2	190	1,050	560	700	1.1	1,100	0.03	- 0.02	
RECo 48/60p	—	R 5-1-3								5.2	48	500	360	600	1.05	1,000	0.05	- 0.3	

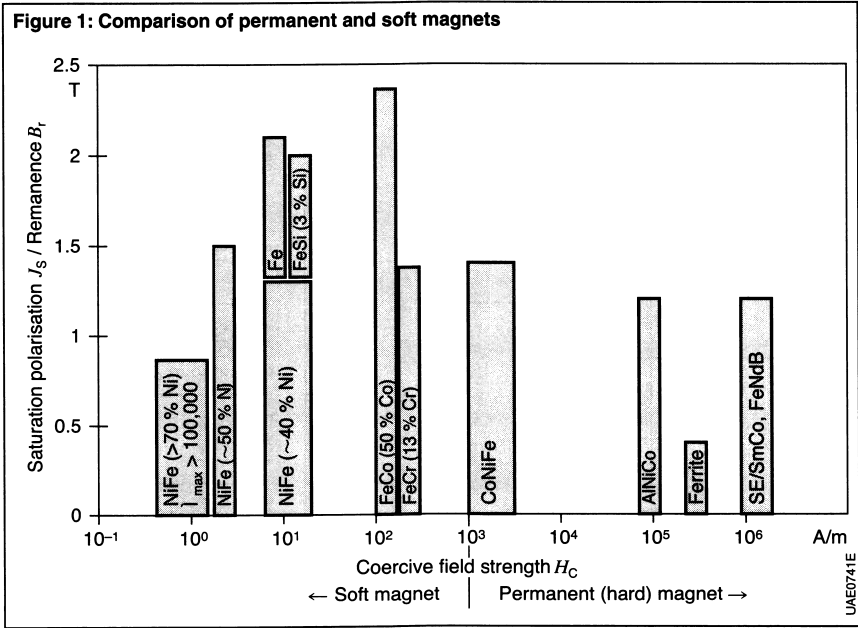
Material		Chemical composition ¹⁾ % by weight										Density $\rho^{(1)}$ g/cm ³	$(BH)^{max(2)}$ kJ/m ³	Rema- nence $B_r^{(2)}$ mT	Coercive field strength ²⁾ of flux density H_{CS} kA/m		Rel. permanent permea- bility ¹⁾ μ_p	Curie temp. ¹⁾ T_c K	Temp. coeff. of polar- ization ¹⁾ $TK (J_s)$ %K	Temp. coeff. of coerc. $TK (H_c)$ %K	Manufacture, processing, applications						
Abbreviated name	Material number	DIN	IEC	Al	Co	Cu	Nb	Ni	Ti	Fe	H_{CS} kA/m				H_{CJ} kA/m	μ_p						T_c K					
CrFeCo 12/4	–	–	R 6-0-1	(no data)							7.6	12	800	40	42	5.5 to 6.5	1,125	–0.03	–0.04								
CrFeCo 28/5	–	–	R 6-1-1								7.6	28	1,000	45	46	3 to 4	1,125	–0.03	–0.04								
REFe 165/170	–	–	R 7-1-1	(no data)							7.4	165	940	700	1,700	1.07	583	–0.1	–0.8								
REFe 220/140	–	–	R 7-1-6								7.4	220	1,090	800	1,400	1.05	583	–0.1	–0.8								
REFe 240/110	–	–	R 7-1-7								7.4	240	1,140	850	1,100	1.05	583	–0.1	–0.8								
REFe 260/80	–	–	R 7-1-8								7.4	260	1,180	750	800	1.05	583	–0.1	–0.8								
Material											Coercive field strength ²⁾ of flux density H_{CS} kA/m	Rel. permanent permea- bility ¹⁾ μ_p	Curie- temp. ¹⁾ T_c K	Temp. coeff. of polariza- tion ¹⁾ $TK (J_s)$ %K	Temp. coeff. of coer. ¹⁾ $TK (H_c)$ %K	Manufacture, processing, applications											
Abbreviated name	Material number	DIN	IEC																								
Ceramic magnets																											
Isotropic																											
Hard ferrite 7/21	1.3641	S 1-0-1									125	210	190	6.5	4.9	1.2	723	–0.2	0.2 to 0.5	Manufacture: sintering. Plastic-bound magnets produced by pressing, injection molding, rolling, extruding. Processing: grinding.							
Hard ferrite 3/18p	1.3614	S 1-2-2									85	175	135	3.2	3.9	1.1											
Anisotropic																											
Hard ferrite 20/19	1.3643	S 1-1-1									170	190	320	20.0	4.8	1.1											
Hard ferrite 20/28	1.3645	S 1-1-2									220	280	320	20.0	4.6	1.1											
Hard ferrite 24/23	1.3647	S 1-1-3									215	230	350	24.0	4.8	1.1	723	–0.2	0.2 to 0.5								
Hard ferrite 25/22	1.3651	S 1-1-5									205	220	370	25.0	4.8	1.1											
Hard ferrite 26/26	–	–	S 1-1-8								230	260	370	26.0	4.7	1.1											
Hard ferrite 32/17	–	–	S 1-1-10								160	165	410	32.0	4.9	1.1											
Hard ferrite 24/35	–	–	S 1-1-14								260	350	360	24.0	4.8	1.1											
Hard ferrite 9/19p	1.3616	S 1-1-3									145	190	220	9.0	3.4	1.1											
Hard ferrite 10/22p	–	–	S 1-2-3								165	225	230	10.0	3.5	1.1											

¹⁾ Standard values. ²⁾ Minimum values. ³⁾ In the range of 273 to 373 K.

References for magnetic materials

[1] DIN IEC 60404-8: Magnetic materials – Part 8: Specifications for individual materials.
8-1: Requirements for individual materials – Hard-magnetic materials (permanent magnets).
8-6: Requirements for individual materials – Soft-magnetic metallic materials.
8-9: Standard requirement for soft magnetic sintered metals.
8-10: Magnetic materials (iron and steel) for relay applications.
[2] DIN 17405: Soft magnetic materials for DC relay; technical terms of delivery.
[3] DIN IEC 60740-2: Laminations for transformers and inductors for use in telecommunication and electronic equipment – Part 2: Specification for the minimum permeabilities of laminations made of soft magnetic metallic materials.
[4] R. Boll: Weichmagnetische Werkstoffe – Einführung in den Magnetismus, VAC Werkstoffe und ihre Anwendungen [Soft magnetic materials – Introduction to magnetism, VAC materials and their applications]. 4th Edition, Publicis Corporate Publishing, 1990.

[5] L. Michalowsky, J. Schneider (ed.): Magnettechnik – Grundlagen, Werkstoffe, Anwendungen [Magnet technology – Foundations, materials, applications]. 3rd Edition, Vulkan-Verlag GmbH, 2006.
[6] Information Sheet 401, “Elektroblech und -band” [Electrical steel sheets and strips], Steel Information Center, Düsseldorf, 2005 edition.
[7] DIN EN 10027-1: Designation systems for steels – Part 1: Steel names.
[8] DIN EN 10106: Cold rolled non-oriented electrical steel sheet and strip delivered in the fully processed state.
[9] DIN EN 10107: Non-oriented electrical steel sheet and strip delivered in the fully processed state.
[10] DIN EN 10341: Cold rolled electrical non-alloy and alloy steel sheet and strip delivered in the semi-processed state.
[11] DIN 41280: Cores of soft magnetic oxides; material properties



Nonmetallic inorganic materials

These materials are characterized by ionic bonds (e.g. ceramic materials), mixed (heteropolar/homopolar) bonds (e.g. glass), or homopolar bonds (e.g. carbon). These kinds of bonds, in turn, are responsible for several characteristic properties: generally poor thermal conductivity and poor electric conductivity (the latter increases with temperature), poor luminous reflectance, brittleness, and thus almost complete unsuitability for cold forming.

Ceramics

Ceramics are at least 30% crystalline in nature; most ceramics also contain amorphous components and pores. Their manufacture is similar to that of sintered metals, however nonmetallic powders or powder mixtures are used; sintering at temperatures generally higher than 1,000 °C gives ceramics their characteristic properties. Ceramic structural parts are sometimes also shaped at high temperatures or even by a melting process, with subsequent crystallization.

Table 1 gives an overview of ceramic materials and their properties.

Glass

Glass is viewed as under-cooled, frozen liquid. Its atoms are only in a short-range order. It is regarded as amorphous. Molten glass turns to solid glass at transformation temperature T_g (T_g is derived from the former designation glass formation temperature). The transformation temperature is dependent on a variety of parameters and therefore not clearly determined (better: transformation range).

Composite materials

Composite materials consist of at least two physically or chemically different components. These components must be tightly bonded together at a specific boundary layer. Under these conditions, it is possible to bond many materials together. The resulting composite material exhibits properties that neither of the original materials has. A distinction is made between the following:

- Particle composite materials: e.g. powder-filled resins, hard metals, plastic-bonded magnets, cermets.
- Laminated composite materials: Such as thermostatic bimetals, sandwich panels.
- Fiber composite materials: Such as glass fiber or carbon fiber-reinforced plastics.

References for Nonmetallic inorganic materials

- [1] DIN EN 623: Advanced technical ceramics; monolithic ceramics; general and textural properties; Part 2: determination of density and porosity.
- [2] DIN EN 843: Advanced technical ceramics – Mechanical properties of monolithic ceramics at room temperature. Part 1: Determination of flexural strength. Part 2: Determination of Young's modulus, shear modulus and Poisson's ratio.
- [3] DIN EN 993: Methods of test for dense shaped refractory products – Part 5: Determination of cold crushing strength.
- [4] DIN EN 821: Advanced technical ceramics – Monolithic ceramics – Thermophysical properties. Part 1: Determination of thermal expansion. Part 2: Determination of thermal diffusivity by the laser flash (or heat pulse) method. Part 3: Determination specific heat capacity.
- [5] DIN EN 60672-1: Ceramic and glass insulating materials – Part 1: Terms and classification.
- [6] www.keramikverband.de.
- [7] www.matweb.com.

Table 1: Ceramic materials

Materials	Composition	$\rho^{(1)}$ g/cm ³	$\alpha_{\text{th}}^{(2)}$ MN/m ²	$\alpha_{\text{th}}^{(3)}$ MN/m ²	$E^{(4)}$ GN/m ²	$\alpha_{\text{th}}^{(5)}$ 10 ⁻⁶ /K	$\lambda^{(6)}$ W/mK	$c^{(7)}$ kJ/kg · K	$\rho_{\text{th}}^{(8)}$ $\Omega \cdot \text{cm}$	$\epsilon_r^{(9)}$	$\tan \rho^{(10)}$ 10 ⁻⁴
Aluminum nitride	AlN > 97%	3.3	250 to 350	1,100	320 to 350	5.1	100 to 220	0.8	> 10 ¹⁴	8.5 to 9.0	3 to 10
Aluminum oxide	Al ₂ O ₃ > 99%	3.9 to 4.0	300 to 400	3,000 to 4,000	380 to 400	7.2 to 8.6	20 to 40	0.8 to 0.9	> 10 ¹¹	8 to 10	2
Aluminum titanate	Al ₂ O ₃ · TiO ₂	3.0 to 3.7	25 to 40	450 to 550	10 to 30	0.5 to 1.5	< 2	0.7	> 10 ¹¹	—	—
Beryllium oxide	BeO > 99%	2.9 to 3.0	250 to 320	1,500	300 to 340	8.5 to 9.0	240 to 280	1.0	> 10 ¹⁴	6.5	3 to 5
Boron carbide	B ₄ C	2.5	300 to 500	2,800	450	5.0	30 to 60	—	10 ⁻¹ to 10 ²	—	—
Cordierite, e.g. C 410, 520 ⁽¹¹⁾	2MgO · 2Al ₂ O ₃ · 5SiO ₂	1.6 to 2.3	5 to 100	300	6 to 60	2.0 to 5.0	1.3 to 2.5	0.8	> 10 ¹¹	5.0	70
Graphite	C > 99.7%	1.5 to 1.8	5 to 30	20 to 50	5 to 15	1.6 to 4.0	100 to 180	—	10 ⁻³	—	—
Porcelain, e.g. C 110 – 120 (non-glazed)	Al ₂ O ₃ 30 to 35% rest SiO ₂ + glass phase	2.2 to 2.4	20 to 100	500 to 550	50	4.0 to 6.5	1.2 to 2.6	0.8	10 ¹¹	6	120
Silicon carbide pressureless- sintered SSiC	SiC > 98%	3.1 to 3.2	400 to 600	> 1,200	400	4.0 to 4.5	90 to 120	0.8	10 ³	—	—
Silicon carbide hot-pressed HPSiC	SiC > 99%	3.1 to 3.2	450 to 800	> 1,500	420	4.0 to 4.5	100 to 120	0.8	10 ³	—	—
Silicon carbide reaction-sintered SiSiC	SiC > 90% + Si	3.0 to 3.1	300 to 400	> 2,200	380	4.2 to 4.8	100 to 160	0.8	10 to 100	—	—
Silicon nitride gas-pressure- sintered GPSN	Si ₃ N ₄ > 90%	3.2	800 to 1,400	> 2,500	300	3.2 to 3.5	30 to 45	0.7	10 ¹²	—	—
Silicon nitride hot-pressed HPSN	Si ₃ N ₄ > 95%	3.2	600 to 900	> 3,000	310	3.2 to 3.5	30 to 45	0.7	10 ¹²	—	—

Materials	Composition	$\rho^{1)}$ g/cm ³	$\sigma_{\text{bs}}^{2)}$ MN/m ²	$\sigma_{\text{dB}}^{3)}$ MN/m ²	$E^{4)}$ GN/m ²	$\alpha^{5)}$ 10 ⁻⁶ /K	$\lambda^{6)}$ W/mK	$c^{7)}$ kJ/kg · K	$\rho_p^{8)}$ $\Omega \cdot \text{cm}$	$\epsilon_r^{9)}$	$\tan \rho^{10)}$ 10 ⁻⁴
Silicon nitride reaction-sintered RBSN	Si ₃ N ₄ > 99%	2.4 to 2.6	200 to 300	< 2,000	140 to 160	2.9 to 3.0	15 to 20	0.7	10 ¹⁴	–	–
Sealtite, e.g. C 220, 221	SiO ₂ 55 to 65 % MgO 25 to 35 % Al ₂ O ₃ 2 to 6 % Alkali oxide < 1.5 %	2.6 to 2.9	120 to 140	850 to 1,000	80 to 100	7.0 to 9.0	2.3 to 2.8	0.7 to 0.9	> 10 ¹¹	6	10 to 20
Titanium carbide	TiC	4.9	–	–	320	7.4	30	–	7 · 10 ⁻⁶	–	–
Titanium nitride	TiN	5.4	–	–	260	9.4	40	–	3 · 10 ⁻⁵	–	–
Titanium dioxide	TiO ₂	3.5 to 3.9	90 to 120	300	–	6.0 to 8.0	3 to 4	0.7 to 0.9	–	40 to 100	8
Zirconium dioxide partially stabilized, PSZ	ZrO ₂ > 90 % rest Y ₂ O ₃	5.7 to 6.0	500 to 1,000	1,800 to 2,100	140 to 210	9.0 to 11.0	2 to 3	0.4	10 ⁸	–	–
Standards for test procedures		DIN EN 623 Part 2 [1]	DIN EN 843 Part 1 [2]	DIN EN 993 Part 5 [3]	DIN EN 843 Part 2 [2]	DIN EN 821 Part 1 [4]	DIN EN 821 Part 2 [4]	DIN EN 821 Part 3 [4]			

The characteristic values for each material can vary widely, depending on the raw material, composition, and manufacturing process. The material data relate to the information provided by various manufacturers.

The designation "KER" corresponds to DIN EN 60-672-1 [5].

Further detailed information on ceramic materials, property tables, applications, etc., can be found on the relevant internet pages, e.g. Informationszentrum Technische Keramik (center of information for technical ceramics) [6], or MatWeb Material Property Data [7]

- 1) Density. 2) Flexural strength. 3) Cold compressive strength. 4) Modulus of elasticity. 5) Coefficient of thermal expansion RT to 1,000 °C.
6) Thermal conductivity at 20 °C. 7) Specific heat. 8) Electrical resistance at 20 °C and 50 Hz. 9) Relative permittivity.
10) Dielectric loss factor at 25 °C and 10 MHz. 11) Properties greatly dependent on porosity set specifically for the application.

Plastics

Plastics are still a relatively “young” material group, the importance of which has continuously grown since the middle of the twentieth century. Products made of plastics play a large role in the various areas of our society today. Thus plastics are used as material in areas such as supply, medical and electrical technology,

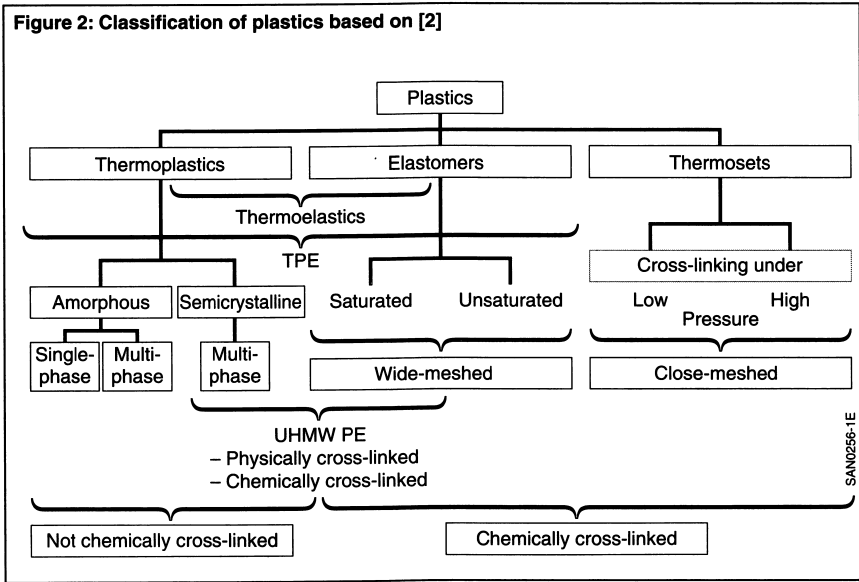
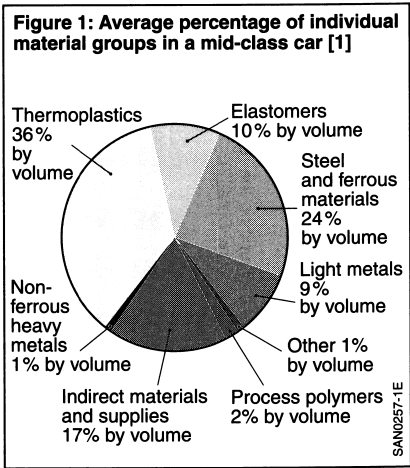
for packaging, household appliances and consumer goods.

Specifically in the automotive industry, plastics have been able to gain increased use more than any other material group (Figure 1). The automotive industry is often the driver behind technological innovations for plastic.

However, the “saturation limit” for using plastics is far from having been reached. We continue to find applications in which conventional materials such as metal are being replaced by plastics.

The increasing use of plastics is mostly due to their less expensive processing options compared to other materials and to the beneficial properties of this “made-to-measure material group”. In addition, new material and process developments are opening new markets and simultaneously offering enormous potential for new, innovative products made of plastic. A strong increase in the plastic content is forecast particularly due to increasing electromobility. The greatest driver is the savings in weight compared to metallic materials.

Plastics is the general umbrella term for polymer materials. A significant characteristic is their macromolecular structure.



Plastics are classified into thermoplastics, thermosets, elastomers and thermoplastic elastomers (Figure 2), which are discussed in more detail in the following.

Thermoplastics

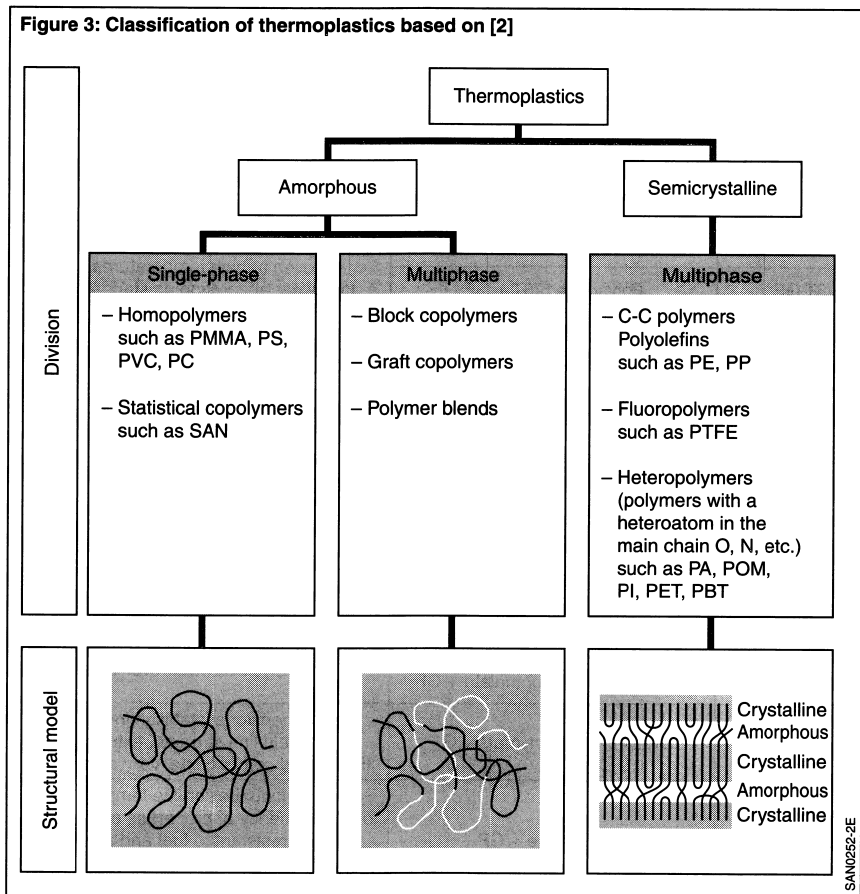
An outstanding characteristic of thermoplastics is the non-cross-linked structure between the macromolecules. This enables repeated plasticity or processability above their service temperature in the melting range. Only thermoplastics can be welded.

The material class of the thermoplastics can be further classified into amorphous thermoplastics and semicrystalline thermoplastics [3]. Semicrystalline ther-

moplastics are composed of amorphous and semicrystalline macromolecular structures. Unlike amorphous thermoplastics, therefore, they exist in multiple phases. Even amorphous thermoplastics can exist in multiple phases if they are modified during their synthesis into the form of a copolymer. The correlations are described in Figure 3 (based on [2]). The block copolymers existing in multiple phases form the link to the material class of thermoplastic elastomers.

The number of manufacturers and commercial types in the area of thermoplastics is very large. Most polymers come with or without different kinds and percentages of fillers and reinforcing materials. Therefore

Figure 3: Classification of thermoplastics based on [2]



the range of properties of thermoplastics available on the market is very large. Table 1 provides an overview of names, codes and example applications of widespread thermoplastics.

Mechanical properties of thermoplastics
Compared to other construction materials, thermoplastics have exceptional viscoelastic and viscoplastic deformation behavior. Consequently, the mechanical behavior is strongly influenced by temperature, load exposure period and loading rate (Figure 4). Amorphous thermoplas-

Table 1: Chemical name and properties of thermoplastics

Code	Chemical name	Description of properties; application examples
ABS	Acrylnitrile-butadien-styrene	High gloss, some transparent types; impact-resistant housing parts
PA 11, 12	Polyamide 11, 12	Tough, hard, and resistant to abrasion, low coefficient of friction, good sound absorption, approx. 1 to 3 % water absorption required for good toughness; PA 11, 12 have much lower water absorption
PA6	Polyamide 6	
PA66	Polyamide 66	
PA6-GF	Polyamide 6 + GF	Impact-resistant machine housings
PA66-GF	Polyamide 66 + GF	
PA6T/6I/66-GF	Polyamide 6T/6I/66 +GF	Rigid machine housings and components even at high temperatures, low water absorption as standard PA
PA6/6T-GF	Polyamide 6/6T + GF	
PAI	Polyamid-imide	Components that are subjected to mechanical or electrical stress, favorable wear characteristics
PBT	Polybutylene terephthalate	Wear-resistant, chemically resistant, hydrolysis in water over 70 °C, very good electrical properties
PBT-GF	Polybutene terephthalate + GF	
PC	Polycarbonate	Tough and rigid over a broad temperature range, components have high rigidity
PC-GF	Polycarbonate + GF	
PE	Polyethylene	Acid-resistant containers and pipes, films
PET	Polyethylene terephthalate	Wear-resistant, chemically resistant, hydrolysis in water over 70 °C
PET-GF	Polyethylene terephthalate + GF	
LCP-GF	Liquid crystal polymers + GF	High dimensional stability under heat, low weld line strength, extremely thin-walled components, very anisotropic
PESU-GF	Polysulfone + GF	High continuous service temperature, low dependency of properties on tempera- ture, not resistant to fuel and alcohol at higher temperatures, dimensionally stable components

Table 1 (continued): Chemical name and properties of thermoplastic plastics

Code	Chemical name	Description of properties; application examples
PEEK-GF	Polyetheretherketone + GF	High-strength components for high temperatures, good antifriction properties, and low wear; very expensive
PMMA	Polymethylmethacrylate	Transparent and in many colors, weatherproof
POM	Polyoxymethylene	Sensitive to acid-induced stress cracking; precision moldings
POM-GF	Polyoxymethylene + GF	
PPE+SB	Polyphenylene ether + SB	Resistant to hot water, flame retardant
PPS-GF	Polyphenylene sulfide + GF	High resistance to heat and media, inherently flame-retarding; underhood components
PP	Polypropylene	Household goods, battery cases, cover hoods, fan impellers (reinforced types)
PP-GF	Polypropylene + GF	
PS	Polystyrene	Transparent and in many colors
PSU-GF	Polysulfone + GF	Low dependence of properties on temperature, not resistant to fuel and alcohol
PVC-P	Polyvinyl chloride (plasticized)	Artificial leather, flexible caps, cable insulation, tubes/hoses, seals
PVC-U	Polyvinyl chloride (unplasticized)	Weatherproof exterior parts, pipes
SAN	Styrene-acrylonitrile	Molded parts with good chemical resistance, also transparent
SB	Styrene-butadiene	Impact-resistant housing parts for many applications
SPS-GF	Syndiotactic polystyrene	Low-warpage, brittle, high mold temperatures required for processing
PI	Polyimide	High resistance to heat and radiation, can be processed only by compacting and sintering
PTFE	Polytetrafluoroethylene	Strong dependence of rigidity on temperature, high resistance to heat, aging and chemicals, can be processed only by compacting and sintering

tics and semicrystalline thermoplastics behave differently. This is due to the different path of the shear modulus curve and the different transition areas associated with this. They are called glass transition temperature T_g and melting range (Figure 5, based on [2] and [3]). With

thermoplastics, the combination of load and higher temperature leads to creep or relaxation due to the viscoelastic and viscoplastic material behavior. In Figure 6, the creep and relaxation behavior of thermoplastics is compared to the material behavior of metals [2].

Figure 4: Depiction of the effect of different test speeds and temperatures on the stress-strain behavior of thermoplastics

- a) Dependency on the test speed,
b) Dependency on the temperature.

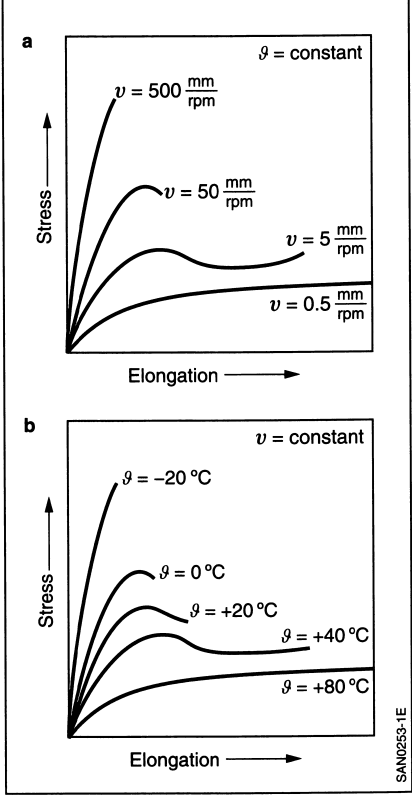


Table 2 and Table 3 show the mechanical properties of amorphous and semicrystalline materials. The values reproduce reference values from literature ([2] and [4]). The values may differ depending on the manufacturer and composition of filler or reinforcing material.

Chemical properties of thermoplastics

The chemical behavior of thermoplastics is determined by the structure of the macromolecules from which they are built. Polar plastics are attacked by polar solvents; non-polar plastics are attacked by non-polar solvents. Substances with a low molecular weight can migrate through solid thermoplastics (permeation).

Permeation of materials with a low molecular weight can trigger formation of stress cracks, much like with metals (referred to here as stress corrosion cracking). Suitable thermoplastics can be selected for a large number of application areas by selecting the right type of thermoplastic, appropriate shaping for the material and article, and optimum production parameters. Table 4 shows the chemical properties of some thermoplastics. The values are reference values from literature ([2], [4], [6]). Since the chemical behavior is difficult to estimate in some cases, consulting the plastics manufacturer or conducting your own measurements is advised.

Durability and aging of thermoplastics

Aging includes all irreversible changes to a material over time [7]. Aging processes change the properties of thermoplastics during a certain period of time. The causes of aging are distinguished between inner (such as residual stress, limited miscibility of additives) and external (energy input from heat and radiation, temperature change, mechanical loading, chemical influences). These causes lead to aging processes that manifest themselves in signs of aging and have a visible or measurable effect on thermoplastics. Examples of this are swelling, aftershrinkage, discoloration and alteration of the mechanical properties, such as brittleness.

Figure 5: Temperature dependence of the dynamic shear modulus

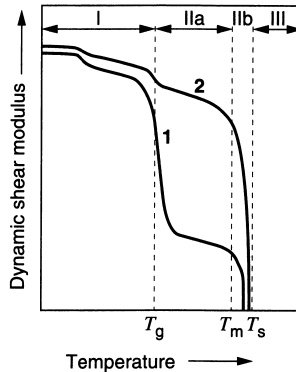
- 1 Amorphous thermoplastics,
 - 2 Semicrystalline thermoplastics.
- T_g Glass transition temperature,
 T_m Melt or flow temperature,
 T_s Crystalline melting temperature.

Amorphous thermoplastics

Area I:
 Glassy state,
 Energy-elastic behavior;
 Scope

Area IIa:
 Entropy-elastic behavior
 (viscoplastic),
 Hot working range

Area III:
 Viscous flow behavior,
 Primary forming
 and welding area



Semicrystalline thermoplastics

Area I:
 Glassy state,
 Energy-elastic behavior,
 Amorphous areas frozen

Area IIa:
 Amorphous parts
 Thermoelastic,
 Semicrystalline
 parts rigid;
 Scope

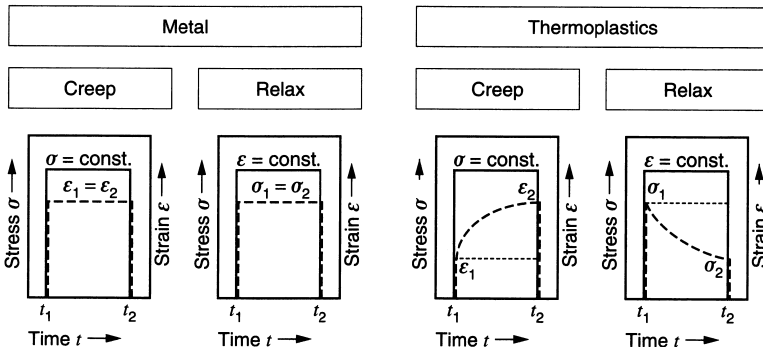
Area IIb:
 Crystallites begin
 to fuse,
 Hot working range

Area III:
 Viscous flow behavior,
 Primary forming
 and welding area

SAN0254-E

Figure 6: Comparison of the influence of time on the mechanical behavior of metals and thermoplastics

T_R Recrystallization temperature, T_g Glass transition temperature.



Influence of t is relatively low
 as long as $T < T_R$

Influence of t is to be noted particularly
 if $T > T_g$
 The smaller the E-modulus,
 the greater the tendency to creep

SAN0255-1E

Table 2: Mechanical properties of amorphous thermoplastics

Code	E module [N/mm ²]	Tensile strain at yield or tensile stress at break [N/mm ²]	Tensile strain at yield or Elongation at fracture [%]	Glass transition temperature [°C]	Continuous service temperature [°C]
ABS	1,300 to 2,700	32 to 45 (y)	15 to 30 (b)	80 to 110	75 to 85
PAI-GF30	10,800	205 (b)	7 (b)	240 to 275	260
PC	2,100 to 2,400	56 to 67 (y)	100 to 130 (b)	150	130
PEI-GF30	9,000	160 (y)	3 (b)	215	170
PESU-GF20	5,700 to 7,500	105 to 130 (b)	2.5 to 3.2 (b)	220 to 225	160 to 200
PMMA	1,600 to 3,600	50 to 77	2 to 10	110	65 to 90
PS	3,200 to 3,250	45 to 65	3 to 4	95 to 100	60 to 80
PSU-GF20	6,200 to 7,000	100 to 115 (b)	2 to 3 (b)	180 to 190	160 to 180
PVC-E	2,000 to 3,000	50 to 60 (y)	10 to 50 (b)	80	65
PVC-S	2,000 to 3,000	50 to 60 (y)	10 to 20 (b)	85	65
SAN	3,600	75 (b)	5 (b)	110	85
PI	3,000* to 3,200*	75 to 100 (b)	n/s	250 to 270	260

* Flexural-E-modulus,
y Value at first maximum value during the tensile test,
b Value at which the specimen breaks,
n/s Not specified.

More physical properties of the thermoplastics

Thermoplastics have good insulation properties against electricity and heat. Compared to other materials, thermoplastics have significantly larger and anisotropic coefficients of thermal expansion ([2], [4], [5]). Some thermoplastics, above all polyamides, absorb a significant amount of water, which results in changes to the mechanical and physical properties as well as the component dimensions.

Processing of thermoplastics

Thermoplastics are usually made available by raw material producers in granular form as sacked goods and can be processed in commercially available injection molding machines and extruders. Since thermoplastics are deformable, they can also be processed using deep-draw or

compression methods. Since thermoplastics are meltable and solidify again, they can be also be welded during the production process and can be recycled to a certain extent.

Raw material producers usually provide recommendations for processing thermoplastics.

Use of thermoplastics

Thermoplastics are used in a wide variety of industries and in very diverse applications. Thermoplastics are used in the packaging industry, construction industry, consumer goods area (household appliances, toys, sports equipment), medical technology and aerospace technology. In the automotive industry, thermoplastics are being used very successfully in the interior, exterior and powertrain areas.

Table 3: Mechanical properties of semicrystalline thermoplastics

Code	E module [N/mm ²]	Tensile strain at yield or tensile stress at break [N/mm ²]	Tensile strain at yield or Elongation at fracture [%]	Glass transition temperature [°C]	Continuous service temperature [°C]
PA 11	1,370 (k)	42 (k, y)	5 (k, y)	49 (tr)	70 to 80
PA 11-GF30	7,300 (k)	134 (k, b)	6 (k, b)	49 (tr)	70 to 80
PA 12	1,600 (tr) / 1,100 (k)	50 (tr, y) / 40 (k, y)	5 (tr, y) / 12 (k, y)	49 (tr)	70 to 80
PA 12-GF30	8,000 (tr) / 7,500 (k)	130 (tr, b) / 120 (k, b)	6 (tr, b) / 6 (k, b)	49 (tr)	70 to 80
PA 6	3,000 (tr) / 1,000 (k)	85 (tr, y) / 40 (k, y)	4.5 (tr, y) / 20 (k, y)	60 (tr)	80 to 100
PA 6-GF30	9,500 (tr) / 6,200 (k)	185 (tr, y) / 115 (k, y)	3.5 (tr, y) / 8 (k, y)	60 (tr)	100 to 130
PA 66	3,000 (tr) / 1,100 (k)	85 (tr, y) / 50 (k, y)	4.4 (tr, y) / 20 (k, y)	70 (tr)	80 to 120
PA 66-GF30	10,000 (tr) / 7,200 (k)	190 (tr, y) / 130 (k, y)	3 (tr, y) / 5 (k, y)	70 (tr)	100 to 130
PAEK-GF30	10,600	168 (b)	2.3 (b)	158	240 to 250
PBT	2,500	60 (y)	3.7 (y) / >50 (b)	60	100
PBT-GF30	10,000	135	2.5 (b)	60	150
PE-LD	200 to 500	8 to 23	300 to 1,000	-30	60 to 75
PET	2,800	80 (y)	4 (y) / 12 (b)	98	100
PET-GF35	14,000	150 (b)	1.5 (b)	98	100
PEEK-GF30	9,700	156 (b)	2 (b)	145	240
POM (H)	3,200	67 to 72 (y)	25–70 (b)	-60	90 to 110
POM (CoP)	2,800	65 to 70 (y)	25–70 (b)	-60	90 to 110
PPS-GF	14,700	195 (b)	1.9 (b)	85 to 95	200 to 240
PP	1,100 to 1,300	30 (y)	20–800 (b)	-10 to 0	100
PTFE	408	25 to 36 (y)	350– 550 (b)	127	260

k Conditioned,

tr Dry,

y Value at first maximum value during the tensile test,

b Value at which the specimen breaks.

Table 4: Chemical properties of thermoplastics

Code	Gasoline	Benzene	Diesel	Alcohol	Mineral oil	Brake fluid	Water cold/hot
ABS	+	-	+	+	+	-	+/+
PA 11	+	+	+	+	+	+	+/O
PA 12	+	+	+	+	+	n/s	+/O
PA6	+	+	+	+	+	+/O	+/O
PA6-GF30	+	+	+	+	+	+/O	+/O
PA66	+	+	+	+	+	+/O	+/O
PA66-GF30	+	+	+	+	+	+/O	+/O
PAI-GF	+	n/s	n/s	+	+	n/s	+/-
PBT	+	O	+	+	+	+	+/-
PBT-GF30	+	O	+	+	+	+	+/-
PC	+	-	O	O	+	n/s	+/O
PC-GF	+	-	O	O	+	n/s	+/O
PE	O	O	+	+	+	n/s	+/+
PET	+	+	+	+	+	n/s	+/-
PET-GF30	+	+	+	+	+	n/s	+/-
PESU-GF	+	-	+	O	+	-	+/O
PEEK-GF	+	n/s	n/s	+	n/s	n/s	+
PI	+	O	+	+	+	n/s	+/O
PMMA	+	O	+	-	+	n/s	+/+
POM (H)	+	+	+	+	+	n/s	+/+
POM-GF (H)	+	+	+	+	+	n/s	+/+
POM (CoP)	+	+	+	+	+	+	+
POM-GF (CoP)	+	+	+	+	+	+	+
PPS-GF40	+	+	+	+	+	+	+/+
PP	O	O	+	+	+	n/s	+/+
PP-GF30	O	O	+	+	+	n/s	+/+
PS	-	-	O	+	O	-	+/+
PSU	+	-	+	O	+	-	+/O
PTFE	+	+	+	+	+	n/s	+/+
PVC-P	-	-	O	-	O	n/s	+/O
PVC-U	+	-	+	+	+	n/s	+/O
SAN	-	-	O	+	+	-	+/+

+ Good resistance

O Limited resistance

- No resistance

n/s Not specified

Thermoplastic elastomers

Thermoplastic elastomers (TPE) form their own plastic material class between thermoplastics and elastomers. They can be processed in a purely physical process that combines high shear forces, heat effects and subsequent cooling (for example, during injection molding or extrusion). Although no chemical cross-linking is required by a time-consuming and high-temperature vulcanization process such as with elastomers, the manufactured parts do in fact have rubber-elastic properties due to their special molecular structure. Renewed influence of heat and shear force causes the material to melt and deform. At the same time, however, this means that the thermoplastic elastomers can withstand far less thermal and dynamic loads than standard elastomers. The thermoplastic elastomers are not a "successor product" of conventional elastomers, but rather an interesting supplement, which combines the processing benefits of thermoplastics with the material properties of the elastomers ([2], [5], [8], [9], [10]).

Types of thermoplastic elastomers

"Thermoplastic elastomers" is the umbrella term for a whole series of different materials. They are always created through blends or block copolymers.

The blends are alloys made of a plastic matrix and a soft elastomer material. Block copolymers are molecular chains with different segments that congregate into hard and soft areas during cooling. Based on DIN EN ISO 18064 [11], the following subdivision can be made.

- TPO: Thermoplastic elastomers on an olefin basis, primarily PP/EPDM, such as Santoprene (Exxon Mobil).
- TPV: Cross-linked thermoplastic elastomers on an olefin basis, primarily PP/EPDM, such as Sarlink (DSM) and Forprene (SoFter).
- TPU: Thermoplastic elastomers on a urethane basis, such as Desmopan (Bayer) and Elastollan (BASF).
- TPC: Thermoplastic copolyester elastomers, such as Hytrel (DuPont) and Riteflex (Ticona).
- TPS: Styrene block copolymers (SBS, SEBS, SEPS, SEEPS and MBS), such as thermoplastic (Kraiburg TPE).
- TPA: Thermoplastic copolyamides, such as PEBAX (Arkema).

Properties of thermoplastic elastomers

The number of diverse commercial types with different properties on the market is very large. They, like thermoplastics, are usually made available by raw material producers in granular form as sacked goods.

Thermoplastic elastomers can be processed very well using injection molding and extrusion processes, since they pass through the plastic, molten state. They can be manufactured in all hardnesses from 5 Shore A to over 70 Shore D. The hardness and the compression set (DVR) are essential characteristics when using thermoplastic elastomers as sealants. Their thermal stability in particular is usually lower than that of elastomers. The maximum continuous service temperatures that can currently be attained for thermoplastic elastomers are approximately 150 °C.

Adhesion to nearly all technical thermoplastics can be achieved through modification. Their flowability as well as density, optics, scratch resistance and other properties can also be adjusted through compounding with a wide variety of fillers and additives.

Application of thermoplastic elastomers

Thermoplastic elastomers find a wide variety of applications in various industries while fulfilling the industry-standard requirements. Thus they are used in the automotive industry for control elements in the interior as well as for window trim in the exterior and for seals close to the engine. Furthermore, they are used in the industrial sector, for example, for tool handles or cable jackets. In the consumer sector, thermoplastic elastomers are found in toys, sports equipment, packaging, and personal care supplies such as toothbrushes and shavers. There are special compounds that satisfy the high requirements even for medical applications. They are used, for example, for drip chambers, seals and medical hoses.

Elastomers

Elastomers (or rubber compounds) are dimensionally stable, but elastically deformable plastics. The essential characteristic of these materials is that they can be stretched to at least twice their length. However, when the tensile load or compressive load is removed, they return to their initial state. This unique ability to reset is also referred to as rubber elasticity.

Elastomers are created from loose, wide-meshed and three-dimensional cross-linking of amorphous preliminary products (rubber). This loose fixation of polymer chains through chemical bonds leads to typical elastic behavior above the glass transition temperature T_g . This value usually lies significantly below 0°C , and therefore below the service temperature of rubber compounds.

On the other hand, if the cross-linking is made solid and close-meshed, we speak of a thermoset. Cross-linked elastomers (just like thermosets) cannot be melted; that is, they decompose at high temperatures without melting.

The starting point for manufacturing elastomer materials is either natural or synthetic rubber. The rubber is mixed with various additives such as fillers, plasticizers, cross-linking chemicals, anti-aging agents and processing aids; then they are cross-linked under the influence of temperature.

During the cross-linking (vulcanization) a chemical reaction is underway, which usually takes place during the molding process under the influence of temperature (150 to 210°C) and pressure. Not until after the vulcanization does the material have its rubber-elastic and mechanical properties, such as the desired hardness, tensile strength and elongation at break ([9], [12], [13], [14], [15]).

Classification of elastomers

Elastomers is an umbrella term for a whole material class whose individual materials can have widely differing properties. They can be classified into groups; based on the DIN ISO 1629 standard [16], the following subdivision can be made.

- R class: Rubbers with an unsaturated carbon chain, such as NR, NBR and SBR.
- M class: Rubbers with a saturated carbon chain, such as ACM, EPDM and FKM.
- O class: Rubbers with carbon and oxygen in the polymer chain, such as ECO.
- U class: Rubbers with carbon, oxygen and nitrogen in the polymer chain, such as polyurethane elastomers AU and EU.
- Q class: Rubbers with silicon and oxygen in the polymer chain, such as silicone elastomers VMQ.

Here, note that the ISO name (for example, NR for Natural Rubber) merely refers to the base rubber.

Properties of elastomers

General basic properties such as thermal and media resistance are already largely determined by the base rubber being used. These can be changed only within certain (very limited) boundaries. Yet within an elastomer class, such as NBR, there is a wide variety of different mixtures that can have significantly different specific properties such as hardness, strength behavior and elasticity. In addition, component suppliers often use their own recipes for their materials and the rubber compounds are often specially designed for the requirements of an application. Unlike thermoplastics or thermoplastic elastomers, non-cross-linked elas-

tomers mixtures are available on the free market only in a few cases and are not offered also in granular form as sacked goods or the like.

Tables 5 and 6 provide an overview of the names, application examples and some important properties of the most common kinds of elastomers. Please note: The data specified there can serve

only as inexact guideline values and must be verified in each individual case of an intended application.

Use of elastomers

Elastomers have the property of being able both to absorb large deformations reversibly and to absorb mechanical energy. Ultimately, this results in diverse

Table 5: Name and application examples of elastomers

Code	Designation	Application examples
M-group		
ACM	Acrylate rubber	Oil circuit (for example, O-rings, radial shaft seal)
AEM	Ethylene acrylate rubber	Oil circuit, dampers, absorbers
EPDM	Ethylene-propylene-diene rubber	Coolant circuit, brake parts, body seals
FFKM	Perfluorinated rubber	Little use (special applications)
FKM	Fluorcarbon rubber	Standard for fuel applications (gas and diesel)
O-group		
ECO	Epichlorhydrin rubber	Fuel hoses, membranes
Q-group		
FVMQ	Fluorosilicone rubber	Membranes, fuel applications
VMQ	Silicone rubber	Turbocharger hoses, exhaust suspension, airbag coating
R-group		
CR	Chloroprene rubber	Bellows, grommets, windshield wipers, V-belts
HNBR	Hydrogenated (acrylo)nitrile-butadiene rubber	Seals in the engine compartment, hoses, drive belts
IIR	Isobutene-isoprene rubber (butyl rubber)	Gas-tight elastomer parts (inner layer of tires), dampers, membranes
NBR	(Acrylo)nitrile-butadiene rubber	Seals, dampers, absorbers, membranes, valves
NR	Natural rubber	(Truck) tires, engine suspensions, chassis mounts
SBR	Styrene butadiene rubber	Passenger car tires, brake parts
U-group		
AU / EU	Polyurethane rubber (polyester / polyether)	Gear wheels, wipers, damping elements (such as foamed material), interior

Table 6: Properties of elastomers

Code	Hardness range in Shore A	Service temperature (continuous) in °C	Resistance to					
			Weather and ozone	Oils (mineral oil, engine oil)	Gasoline	Diesel fuel	Water	Brake fluid (glycol-based)
M-group								
ACM	50 to 90	– 25 to +150 ¹⁾	1–2	1	3–4	3	4	4
AEM	50 to 90	– 35 to +150	1	1–2	3–4	3	3	4
EPDM	30 to 95	– 50 to +125 ²⁾ – 50 to +150 ³⁾	1	4	4	4	1	1
FFKM	60 to 90	– 15 to +260	1	1	1	1	1	1
FKM	55 to 90	– 20 to +200 ¹⁾	1	1	1	1	2	4
O-group								
ECO	50 to 90	– 40 to +120	1–2	2	2	2	3	4
Q-group								
FVMQ	30 to 80	– 55 to +175	1	1	2	1–2	1–2	4
VMQ	20 to 80	– 60 to +200	1	2–3	4	3	1–2	2–3
R-group								
CR	30 to 90	– 40 to +110	2–3	2–3	3–4	3	2–3	3
HNBR	40 to 90	– 30 to +130 ⁴⁾	2	1	2	1	1	2
IIR	40 to 85	– 40 to +120	2–3	4	4	4	1	1
NBR	35 to 95	– 30 to +100 ⁴⁾	3–4 ⁶⁾	1	2	1	1–2	3–4
NR	30 to 95	– 55 to +80	4	4	4	4	1	1
SBR	30 to 95	– 50 to +100	4	4	4	4	1	1
U-group								
AU / EU	50 to 98	– 40 to +90 ⁵⁾	2	1–2	3	2–3	4	4

1 Very good resistance (no or little effect)

2 Good resistance (moderate effect)

3 Limited resistance (significant effect)

4 No resistance (unsuitable)

¹⁾ Special materials with better cold resistance are possible

²⁾ Sulfur cross-linked

³⁾ Peroxide cross-linked

⁴⁾ Cold resistance, depending on the composition of the polymer

⁵⁾ Materials with better resistance to heat and hydrolysis are possible

⁶⁾ Materials with better ozone resistance are possible

applications for this material class. Thus rubber compounds are used to manufacture products that bridge tolerances, permit movement between various components, constitute static and dynamic seals, diminish vibrations and act as springs.

Typical technical applications include tires, seals (see also Elastomer materials, Seals), V-belts, hoses, flexible couplings, cable jackets, bearing elements, retaining elements, shock absorbers, windshield wipers, conveyor belts, roofing foils and shoe soles. But elastomers are also used to manufacture items for everyday use such as rubber boots, erasers, rubber bands, balloons, condoms, rubber gloves, pacifiers and wet suits.

Thermosets

Thermosets are the first plastics to have been manufactured for industrial production using a synthetic method. As early as 1910, Bakelite, a resin made of formaldehyde and phenol, came into use. In the 1920s and 1930s, molding resin masses with urea and melamine jointed it on the market. Polyester and epoxy resins were used for the first time after the end of World War II [17].

In general, thermosets refer to plastics with close-meshed molecular chains that are tightly cross-linked to each other. This results in the typical properties of thermosets:

- High strengths and rigidities with simultaneously low density,
- High dimensional stability under heat and thermal resistance,
- Good resistance to chemicals,
- High brittleness.

The varied types of thermosets find application in a wide variety of areas in the automotive industry. In addition to structural components, which will be covered in more detail below, this includes paint and adhesive systems, sealing compounds and substrates for interconnect devices in electronics.

The good thermal properties of thermosets make them particularly suitable for use in thermally stressed areas. This specifically relates to applications in the engine compartment of automobiles. Typical examples of components made of thermoset include water pump housings, belt pulleys and impeller wheels [18]. New developments in the area of tribologically modified PF molding compounds also open up further use options, such as for bearing elements, sliding elements and guide elements [19].

Due to the large number of thermosetting resin systems, the section below only covers the technically relevant materials (Table 7).

Unsaturated polyester resins

Unsaturated polyester resins (UP resins) fully cure through radical polymerization under the influence of heat. They can be adjusted with almost no shrinkage, which leads to very good dimensional stability. This is why components with special requirements for precision, such as headlamp reflectors, are manufactured from these kinds of materials. Furthermore, components made of UP resins feature good electrical properties, which helped them find increased usage earlier in ignition systems of motor vehicles (ignition distributors).

Epoxy resins

Epoxy resins (EP resins) cure in a polymerization reaction and therefore, unlike the phenol-formaldehyde resins, do not split off any volatile reaction products. Compared to other thermosetting materials, EP resins have a particularly low viscosity, which simplifies the procedure for processing them. Due to the high material prices, however, use of epoxides is limited. EP resins are used, for example, for encapsulating electronic components (special low-pressure EP is used for this). Moreover, EP resin is increasingly being used in continuous-fiber-reinforced (mostly carbon fiber-reinforced) structural lightweight components as matrix materials in the automotive industry.

Table 7: Thermosets

Type	Type of resin	Filler	$t_g^{(1)}$ °C	$\sigma_{bs}^{(2)}$ min. N/mm ²	$\alpha_n^{(3)}$ min. kJ/m ²	CTI ⁽⁴⁾ min. grade	Properties, application examples
Thermosets (new standards)							
– Pourable phenol molding compounds (PF–PMC) DIN EN ISO 14526 [33] – Pourable melamine-formaldehyde molding compounds (MF–PMC) DIN EN ISO 14528 [34] – Pourable melamine/phenolic molding compounds (MP–PMC) DIN EN ISO 14529 [35] – Pourable unsaturated polyester molding compounds (UP–PMC) DIN EN ISO 14530 [36] – Pourable epoxy-resin molding compounds (EP–PMC) DIN EN ISO 15252 [37]							
(WD30 + MD20) to (WD40 + MD10) (31 and 31.5) ⁷⁾	Phenol ⁸⁾	Wood flour	160/140	70	6	CTI 125	For parts subject to high electrical loads. ⁹⁾
(LF20 + MD25) to (LF30 + MD15) (51) ⁷⁾		Cellulose ⁵⁾	160/140	60	5	CTI 150	For parts with good insulating properties in low-voltage range. Type 74 impact-resistant. ⁹⁾
*SS40 to SS50 (74) ⁷⁾		Cotton fabric shreds ⁵⁾	160/140	60	12	CTI 150	
(LF20 + MD25) to (LF40 + MD05) (83) ⁷⁾		Cotton fibers ⁶⁾	160/140	60	5	CTI 150	Tougher than type 31. ⁹⁾
–		Glass fibers, short	220/180	200	12	CTI 125	High mechanical strength. Very good resistance to automotive fluids, low swelling.
–		Glass fibers, long	220/180	230	17	CTI 175	
–		Carbon fibers	220/180	250	14	–	High rigidity, low density, good wear properties. Not suitable for electrical applications (conductive).
(WD30 + MD15) to (WD40 + MD05) (150) ⁷⁾	Melamine	Wood flour	160/140	70	6	CTI 600	Resistant to glow heat, superior electrical properties, high shrinkage factor.

Table 7 (continued): Thermosets

Type	Type of resin	Filler	$t_g^{(1)}$ °C	$\sigma_{BS}^{(2)}$ min. N/mm ²	$\alpha_n^{(3)}$ min. kJ/m ²	CTI ⁽⁴⁾ min. grade	Properties, application examples
Thermosets							
LD35 to LD45 (181) ⁷⁾	Melamine-phenol	Cellulose	160/140	80	7	CTI 250	For parts subject to electrical and mechanical loads.
(GF10 + MD60) to (GF20 + MD50) (802 and 804) ⁷⁾	Polyester	Glass fibers, inorganic fillers	220/170	55	4.5	CTI 600	Types 801, 804: low molding pressure required (large-area parts possible); types 803, 804: glow-heat resistant.
MD65 to MD75	Epoxy	Rock flour	200/170	80	5	CTI 600	Very good dielectrical properties. Sheathing sensors and actuators.
(GF25 + MD45) to (GF35 + MD35)		Glass fibers/ mineral	230/190	160	10	CTI 250	
–	Epoxy low-pressure molding compounds	SiO ₂ (spherical)	250/200	120	6	–	Chip encapsulation (thin-bond wires).

¹⁾ Maximum service temperature, short term (100 h)/continuous (20,000 h). ²⁾ Flexural strength. ³⁾ Impact strength (Charpy).

⁴⁾ Tracking resistance according to DIN IEC 112, Comparative Tracking Index (CTI).

⁵⁾ With or without addition of other organic fillers. ⁶⁾ And/or wood flour. ⁷⁾ Old designation in parentheses.

⁸⁾ Do not use types 13 to 83 (purely organically filled compounds) for new applications (availability no longer guaranteed).

⁹⁾ Barely used any more in new applications. Supply not guaranteed in the long term.

Phenol-formaldehyde resins

Phenol-formaldehyde resins (PF resins) are manufactured from phenol and formaldehyde using polycondensation. They have high mechanical rigidities and strengths and feature high chemical resistance, thermal resistance and dimensional stability under heat. In addition, there are tribologically modified forms that, however, cost significantly more because carbon fibers are added. PF resins are inherently flame retardant, which makes them ideal for applications in engine compartments.

Extenders

Filler are classified in accordance with DIN EN ISO 1043-2 [20]. A distinction is made here between material (such as carbon, glass, mineral), the form and structure (such as fibers, balls, powder) and special properties (such as flame retardant, thermal resistance) [4]. Technically relevant fillers primarily include glass fibers and glass balls as well as mineral fillers. Particularly in the aerospace industry, but also increasingly in the automotive industry, carbon fibers are also being used as reinforcement materials for thermosets.

Processing method

Thermosetting molding compounds are primarily processed using compacting or injection molding. There is also a wide variety of other processes in which the listed variants are combined or further developed.

Compacting

For compacting, pressure and temperature simultaneously act upon the molding compound. The mass is fed into the mold either without a form or in tablet form. Pressure brings the mass into the desired form, while the temperature cross-links the material. Due to the low machine-based effort, compacting is the least expensive processing method. However, this method can be used to manufacture only simple, large-format components. Above all, the manufactured components feature low level of orientation and the need for large amounts of finishing.

Injection molding

Here it is necessary to distinguish between easy-flowing granulate material and polyester bulk molding compound. Exceptionally high productivity is achieved when processing easy-flowing granulate material where the length of cycle times is largely dependent on component thickness. The barrel temperatures during injection molding are between 80 and 100°C, while the mold temperatures range from 160 to 190°C. While the injected mass is curing in the mold, new material for the next cycle is already being plasticized in the screw.

When processing polyester bulk molding compound (UP BMC), it is important to note that a stuffing unit is required in addition to the conventional machine equipment to convey the material into the screw. Then the material is homogenized in the injection barrel. Melting is not necessary; the barrel temperatures are usually around 25 to 35°C.

Insulating materials

Electrical insulation plays a critical role in the proper functioning and service life of items such as alternators, engines and electrical devices, not just in motor vehicles.

Unfilled polymers have the best electrical insulation properties. Each addition of fillers reduces the dielectric strength of a polymer material by forming interfaces between the filler and polymer matrix as well as by excessive stress due to different dielectric properties.

The electrical insulation in motor vehicles not only has to reliably prevent disruptive electrical breakdowns, it also has to dissipate heat losses that arise, absorb mechanical loads and be resistant to liquids commonly used in motor vehicles. These properties typically have to be ensured in a temperature range from -40 to +180°C. For this reason, a system of several materials is usually used as an insulating system; a single material is rarely used.

As an example here we will discuss surface insulators. These are usually used as more or less flexible combined insulation materials in electrical machines such as starters, alternators, hybrid engines and traction motors.

Plastic foils are flexible combined insulation materials that are bonded to pressboard, nonwoven materials or paper made of polymers. Flexible combined insulation materials often have a three-ply design with plastic foil as the middle layer. Depending on the combination of materials, they have different continuous service temperatures, tensile strengths and elongations at break, dielectric strengths, rigidities and impregnabilities.

The combination of plastic foils with fiber or nonwoven materials has technical advantages. Plastic foil made of unfilled polymers provides excellent electrical properties. The fiber or nonwoven materials, by contrast, have good impregnability and protect the foil against mechanical and thermal loads.

Foils made of polyester or polyimide are predominantly used as initial components as well as fiber or nonwoven materials made of organic fibers, polyester or aramide.

Flexible combined insulation materials are standardized as DIN EN 60626-1 ([21], definitions, general requirements), DIN EN 60626-2 ([21], test methods) and DIN EN 60626-3 ([21], properties of individual combinations of materials). In addition, there are standards for other surface insulators.

- Insulating foils: DIN EN 60674-1 ([22], definitions, general requirements), DIN EN 60674-2 ([22], test methods), DIN EN 60674-3-1 to -3-8 ([23], [24], properties of individual materials).
- Panel and roller pressboard: DIN EN 60641-1 ([25], definitions, general requirements), DIN EN 60641-2 ([26], test methods), DIN EN 60641-3-1, -3-2 ([26], properties of individual materials).
- Laminates: DIN EN 60893-1 ([27], definitions, general requirements), DIN EN 60893-2 ([27], test methods), DIN EN 60893-3-1 to -3-7 ([27], properties).

Sealing compounds

Sealing compounds (or cast resins) refer to reactive synthetic resins that are processed as a liquid for the finished product and solidify as this or a component part of it. While still a liquid, the resin is poured into a reusable or disposable mold. This results in either pure cast resin bodies with free form surfaces or other parts are included ([28], [29], [30], [31], [32]).

Unlike with meltable plastics (thermo-plastics), the solidification occurs through a chemical cross-linking reaction and is irreversible (thermoset). They are usually poured in

- To cover and protect parts against penetration of moisture, dust, foreign matter, water, etc.
- To fix parts into place with respect to each other
- To increase the mechanical stability and the vibration and shock resistance
- To provide electrical insulation, that is, to increase the dielectric strength and contact protection
- To dissipate any heat losses that arise.

The range of applications for sealing compounds is broad and multi-faceted. As a result, the functions and requirements are also multi-faceted, ranging from processing and curing to the properties of the later application area. Selection criteria include:

- For application of liquid, not yet hardened sealing compound: viscosity, pot life, available systems engineering (manual, metering or molding system).
- The requirements that are placed on the cured molding compound in later use, such as hardness, elasticity, stretchability or pliability, thermal, mechanical, electrical and chemical loads of the molded components.

Due to their mechanical and electrical properties (Table 8), sealing compounds are ideal for use in electrical engineering and electronics. Typical applications of sealing compounds include:

- Molding and manufacturing electro-technical components (such as ignition coils, transformers, insulators, capacitors, semiconductors, assemblies)
- Molding of open contact points for cables and lines

Different material classes are used depending on the requirements profile (such as service temperatures, chemical resistance, mold geometry). The sealing compounds used most frequently in electronics potting belong to the epoxide, polyurethane and silicone material classes. Less widespread, for example, are polyester and hot melt sealing compounds (hot melts, hot melt adhesives).

Table 8: Properties of different sealing compounds (examples)

Properties	Standard	Unit	Epoxy resin, unfilled	Epoxy resin, filled (quartz powder)	Unsaturated polyester resins	Polyurethanes	Silicone
Tensile strength	ISO 527-1 [38]	N/mm ²	60 to 90	80 to 100	30 to 80	3 to 80	0.3 to 10
Elongation at break	ISO 527-1 [38]	%	3 to 8	0.8 to 1.1	2 to 4	0.5 to 80	50 to 700
Flexural strength	ISO 178 [39]	N/mm ²	60 to 140	110 to 120	60 to 140	40 to 140	n/a
Impact strength	ISO 179-1 [40]	kJ/m ²	15 to 30	10 to 12	8 to 15	40	n/a
Glass transition temperature T_G		°C	+ 70 to + 200	+ 70 to + 200	+ 70 to + 150	– 40 to + 130	– 120
E module		N/mm ²	2,000 to 4,000	5,000 to 8,000	3,500	250 to 3,000	0.005 to 5
Continuous temperature resistance	DIN EN 60216-1 [41]	°C	110 to 200	110 to 200	120 to 140	90 to 140	150 to 250
Specific volume resistance		Ωcm	10 ¹⁴ to 10 ¹⁶	10 ¹⁴ to 10 ¹⁶	10 ¹³	10 ¹³ to 10 ¹⁶	10 ¹⁵ to 10 ¹⁷
Exothermy during curing			High	Medium	Very high	Very low	Very low
Coefficient of thermal expansion		K ^{–1} 10 ^{–6}	80 to 100	30 to 70	60 to 80	50 to 150	300
Shrinkage		%	0.5 to 2	0.1 to 1	3 to 9	0.2 to 1	0.1 to 2

+ = Good resistance
O = Limited resistance
– = No resistance
n/a = Not applicable

There are also extensive means of using chemicals and formulas to adapt the sealing compounds to the typical requirements for an application:

- Variations in chemicals provide different resin and curing systems
- Variations in formulas provide fillers, pigments, accelerators, toughness modifiers, wetting agents, degassing agents and anti-sedimentation agents

The chemical resistance of sealing compounds depends on the components of the cast resin, the cross-link density and the degree of cross-linking. As a rule of thumb, hard sealing compounds are more resistant than soft ones.

The final properties of the cured sealing compounds largely depend on the kind and quantity of ingredients used (hardeners, diluents, fillers) and the curing conditions. Therefore generally applicable specifications about physical properties are not possible.

Epoxide systems

Epoxides have been widely used for many years. They are generally hard and capable of carrying loads, their volume shrinkage during curing is rather low.

Characteristics include their excellent mechanical properties, good temperature tolerance, good adhesive strength on a wide variety of surfaces, and likewise good chemical resistance. The cross-linking or curing process is generally slow, particularly if only small volumes react with each other. More reactive hardeners could be used, but that can cause a substantially stronger exothermic reaction, creating stress for components and the printed circuit board.

Fillers are an essential component of epoxy resin formulations. They reduce costs and improve mechanical properties. At the same time, they lead to an increase in the crack resistance and rigidity, as well as to a reduction in thermal shrinkage, lowering lower inner stresses. Quartz powder is a standard filler in electronic applications.

Epoxide systems are frequently used, for example, as a thermally conductive sealing compound for all kinds of solenoids.

Polyurethane systems

Even after curing, polyurethane sealing compounds can still expand and flex (even if only somewhat), which is particularly important when sensitive components (such as ferrite on printed circuit boards) are cast.

The curing reaction for polyurethane systems is less exothermic than for epoxide sealing compounds. The volume shrinkage after the hardening is low and a broad spectrum of hardness and/or elasticity is available. The chemical and mechanical resistances are good.

Silicone

Silicone-based sealing compounds are usually significantly more expensive than epoxide or polyurethane sealing compounds, but they are used where typical continuous service temperatures are above 180°C. The exothermic heat buildup during the hardening is also only very low.

Polyester systems

These systems have very good resistances to a wide variety of media, but a very strong heat buildup during hardening and a high degree of shrinkage after the hardening as a result. Under certain circumstances this leads to thermomechanical damage to the cast components, even to the point of having components tear away from the substrate.

References for plastics

- [1] M. Gehde, S. Englich, G. Hülde, M. Höer: Schlummerndes Potenzial für den Leichtbau [The dormant potential for lightweight construction]. In: *Plastverarbeiter* (10), p. 80-83, 2012.
<http://www.plastverarbeiter.de/36471/schlummerndes-potenzial-fuer-den-leichtbau/>
- [2] P. Elsner, P. Eyerer, T. Hirth: Domininghaus – Kunststoffe: Eigenschaften und Anwendungen [Plastics: Properties and applications]. 8. Aufl., Springer, 2012.
- [3] DIN 7724:1993: Polymere Werkstoffe – Gruppierung polymerer Werkstoffe aufgrund ihres mechanischen Verhaltens [Polymer materials – Grouping polymers by their mechanical characteristics].
- [4] W. Hellerich, G. Harsch, S. Haenle: Werkstoff-Führer Kunststoffe – Eigenschaften, Prüfungen, Kennwerte [Plastics material guide – Properties, Testing and Characteristics]. 10. Aufl., Carl Hanser Verlag, 2010.
- [5] E. Baur, S. Brinkmann, N. Rudolph, T.A. Osswald, E. Schmachtenberg: Saechtling Kunststoff Taschenbuch [Saechtling plastics pocketbook], 31. Aufl., Carl Hanser Verlag, 2013.
- [6] G.W. Ehrenstein, S. Pongratz: Beständigkeit von Kunststoffen [The durability of plastics]. 1. Aufl., Carl Hanser Verlag, 2007.
- [7] DIN 50035:2012: Begriffe auf dem Gebiet der Alterung von Materialien – Polymere Werkstoffe [Concepts in the field of material aging – Polymers].
- [8] T. Dolansky, M. Gehringer, H. Neumeier: TPE-Fibel – Grundlagen, Spritzguss [TPE primer – Principles, casting]. 1. Aufl., Dr. Gupta Verlag, 2007.
- [9] F. Röthemeyer, F. Sommer: Kautschuk Technologie: Werkstoffe – Verarbeitung – Produkte [Rubber technology: Materials – Processing – Products]. 3. Aufl., Carl Hanser Verlag, 2013.
- [10] G. Holden, H.R. Kricheldorf, R. Quirk: Thermoplastic Elastomers [Thermoplastic elastomers]. 3. Aufl., Hanser Gardner Publ., 2004.
- [11] DIN EN ISO 18064: Thermoplastic elastomers – Nomenclature and abbreviated terms.
- [12] K. Nagdi: Gummi-Werkstoffe [Rubber materials]. 3. Aufl., Dr. Gupta Verlag, 2004.
- [13] W. Hofmann, H. Gupta: Handbuch der Kautschuk-Technologie [Handbook of rubber technology]. Dr. Gupta Verlag, 2001.
- [14] J. Schnetger: Lexikon Kautschuk-technik [Glossary of rubber technology]. 3. Aufl., Beuth, 2004.
- [15] G. Abts: Einführung in die Kautschuk-technologie [Introduction to rubber technology]. Carl Hanser Verlag, 2007.
- [16] DIN ISO 1629: Rubber and latices – Nomenclature.
- [17] G.W. Becker, D. Braun, W. Woebcken (Hrsg.): Kunststoff Handbuch – Band 10: Duroplaste [Handbook of Plastics – Volume 10: Thermosets]. 2. Aufl., Hanser Fachbuch, 1988.
- [18] E. Bittmann: Duroplaste kommen ins Rollen [Thermosets get rolling]. In: *Kunststoffe* 3/2003, A25-A27.
<https://www.kunststoffe.de/kunststoffe-zeitschrift/archiv/artikel/automobilanwendungenduroplaste-kommen-ins-rollen-530048.html>.
- [19] E. Bittmann: Duroplaste [Thermosets]. In: *Kunststoffe* 10/2005, 168–172.
<https://www.kunststoffe.de/kunststoffe-zeitschrift/archiv/artikel/hoffnungsvolle-auftragslage-duroplaste-und-fvk-537533.html>.
- [20] DIN EN ISO 1043-2: Plastics – Symbols and abbreviated terms – Part 2: Fillers and reinforcing materials. [21] DIN EN 60626: Combined flexible materials for electrical insulation. Part 1: Definitions and general requirements. Part 2: Methods of test. Part 3: Specifications for individual materials.
- [22] DIN EN 60674: Plastic films for electrical purposes. Part 1: Definitions and general requirements. Part 2: Methods of test.
- [23] DIN EN 60674-3-1: Plastic films for electrical purposes – Part 3: Specifications for individual materials. Sheet 1: Biaxially oriented polypropylene (PP) film for capacitors.
- [24] DIN EN 60674-3-2: Specification for plastic films for electrical purposes – Part 3: Specifications for individual materials. Sheet 2: Requirements for balanced biaxially oriented polyethylene terephthalate (PET) films used for electrical insulation.

Sheet 3: Requirements for polycarbonate (PC) films used for electrical insulation.

Sheets 4 to 6: Requirements for polyimide films used for electrical insulation.

Sheet 7: Requirements for fluoroethylene-propylene (FEP) films used for electrical insulation.

Sheet 8: Balanced biaxially oriented polyethylene naphthalate (PEN) films used for electrical insulation.

[25] DIN EN 60641-1: Specification for pressboard and presspaper for electrical purposes.

Part 1: Definitions and general requirements.

[26] DIN EN 60641-2: Pressboard and presspaper for electrical purposes.

Part 2: Methods of test.

Part 3: Specifications for individual materials.

Sheet 1: Requirements for pressboard.

Sheet 2: Requirements for presspaper.

[27] DIN EN 60893-1: Insulating materials – Industrial rigid laminated sheets based on thermosetting resins for electrical purposes.

Part 1: Definitions, designations and general requirements.

Part 2: Methods of test.

Part 3-1: Specifications for individual materials – Types of industrial rigid laminated sheets.

Part 3-2: Specifications for individual materials – Requirements for rigid laminated sheets based on epoxy resins.

Part 3-3: Specifications for individual materials – Requirements for rigid laminated sheets based on melamine resins.

Part 3-4: Specifications for individual materials – Requirements for rigid laminated sheets based on phenol resins.

Part 3-5: Specifications for individual materials – Requirements for rigid laminated sheets based on polyester resins.

Part 3-6: Specifications for individual materials – Requirements for rigid laminated sheets based on silicone resins.

Part 3-7: Specifications for individual materials – Requirements for rigid laminated sheets based on polyimide resins.

[28] R. Stierli: Epoxid-Gieß- und Imprägnierharze für die Elektroindustrie [Epoxide casting and impregnating resins for the electrical industry]. In: Wilbrand Woebcken (Hrsg.): Duroplaste – Kunststoff-Handbuch

[Thermosets – The plastics handbook], Band 10. 2. Aufl., Hanser Fachbuch, 1988.

[29] W. Becker, D. Braun, G. Oertel (Hrsg.): Polyurethane – Kunststoff-Handbuch [Polyurethane – The plastics handbook], Band 7. 3. Aufl., Hanser Fachbuch, 1993.

[30] Dr. Werner Hollstein, Huntsman Advanced Materials GmbH: Einführung in die Chemie der Epoxidharze und Formulierungskomponenten [Introduction to the chemistry of epoxide resins and formulation components].

[31] <http://www.electrolube.com>.

[32] Lackwerke Peters GmbH & Co. KG: Technische Informationen TI 15/2.

<https://www.peters.de/de/download-center>.

[33] DIN EN ISO 14526-1: Plastics – Phenolic power moulding compounds (PF-PMCs) – Part 1: Designation system and basis for specifications.

[34] DIN EN ISO 14528-1: Plastics – Melamine-formaldehyde powder moulding compounds (MF-PMCs) – Part 1: Designation system and basis for specifications.

[35] DIN EN ISO 14529-1: Plastics – Melamine/phenolic power moulding compounds (MP-PMCs) – Part 1: Designation system and basis for specifications.

[36] DIN EN ISO 14530-1: Plastics – Unsaturated-polyester power moulding compounds (UP-PMCs) – Part 1: Designation system and basis for specifications.

[37] DIN EN ISO 15252-1: Plastics – Epoxy power moulding compounds (EP-PMCs) – Part 1: Designation system and basis for specifications.

[38] ISO 527-1: Plastics – Determination of tensile properties – Part 1: General principles.

[39] ISO 178: Plastics – Determination of flexural properties.

[40] ISO 179-1: Plastics – Determination of Charpy impact properties – Part 1: Non-instrumented impact test.

[41] DIN EN 60216-1: Electrical insulating materials – Thermal endurance properties – Part 1: Ageing procedures and evaluation of test results.

Heat treatment of metallic materials

Hardness

Hardness is a property of solid materials that defines the resistance of a materials against penetration by a harder solid body. In metallic materials the hardness is used to assess mechanical properties such as strength, machinability, malleability, and resistance to wear. DIN EN ISO 18265 [1] defines guidelines for converting hardness to tensile strength.

Hardness testing

Hardness testing is a nondestructive way of obtaining information about the mechanical properties of a material in a relatively short time.

The test data are generally derived from the size or depth of the deformation produced when a specified indenter is applied at a defined pressure. A distinction is made between static and dynamic testing. Static testing is based on measurement of the permanent impression left by the indenter. Conventional hardness tests include the Rockwell, Vickers and Brinell procedures. Figure 1 compares the fields of application of hardness testing based on these different procedures.

Dynamic testing monitors the rebound height of a test tool accelerated against the surface of the test specimen.

Another option for obtaining an index of surface hardness is to scratch the surface with a harder test tool and then measure the groove width.

Hardness-testing methods

Rockwell hardness (DIN EN ISO 6508, [2])
This method is particularly suitable for fast, automated testing of metallic workpieces, but places specific demands on clamping the workpiece in the test equipment. It is unsuitable for workpieces which as a result of their geometry give in the test equipment (e.g. pipes).

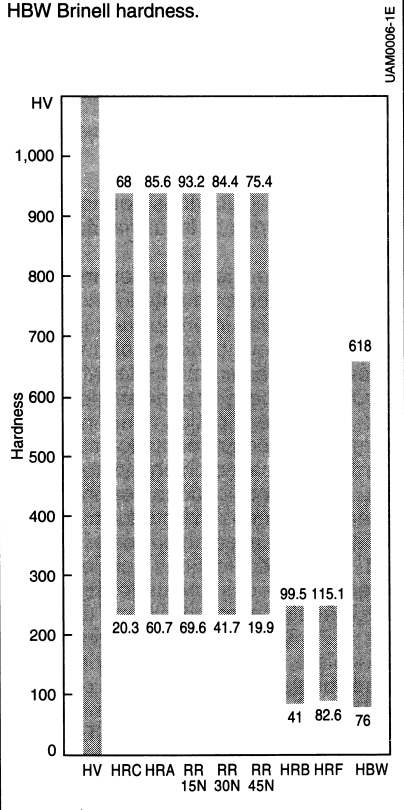
This method involves a test tool (indenter) of defined size, shape (tapered or spherical) and material (steel, hard metal or diamond) being pressed in two stages into the test specimen. In this process, after the preliminary test force is applied,

the additional test force is exerted for a defined period. Initially, the factor e is calculated from the remaining penetration depth h for the HRA, HRB, HRC and HRF methods:

$$e = \frac{h}{0.002}.$$

Figure 1: Comparison of hardness ranges for different test methods for unalloyed and low-alloy steels and cast steel

The figures at the range extremities indicate the hardness data for the respective method. HV Vickers hardness, HR Rockwell hardness, HBW Brinell hardness.



For the HR..N and HR..T methods the factor e is calculated from the penetration depth h with the equation:

$$e = \frac{h}{0.001}.$$

With the factor e the Rockwell hardness HR can now be calculated for the HRA, HRC, HR..N and HR..T methods:

$$HR = 100 - e.$$

If the HRB and HRF methods are used, the Rockwell hardness HR is calculated according to the of following equation:

$$HR = 130 - e.$$

The test specimen's surface should be smooth and as flat as possible. When testing on convex-cylindrical or spherical surfaces, the value determined must be corrected as a function of the hardness.

The abbreviation for the selected test method should be appended to the numerical value when specifying hardness (e.g.: 65 HRC, 76 HR45N). The designations provide an indication of the indenter used (diamond cone or ball), the preliminary test force and the total test force.

Depending on the indenter used and the total test force applied, different hardness scales with the abbreviations HRA, HRB, HRC, HRD, HRE, HRF, HRG, HRH, HRK, HR15N, HR30N, HR45N, HR15T, HR30T and HR45T are used.

Advantages of the Rockwell test method include minimal specimen preparation and rapid measurement. This test process can also be fully automated. Any tester vibration, or shifting and movement of either the test specimen or the indenter can lead to testing errors, as can an uneven support surface or a damaged indenter.

Brinell hardness (DIN EN ISO 6506, [3])

This method is used for metallic materials of low to medium hardness. The test tool (indenter) is a hard-metal ball with the diameter D . It is pressed with test force F vertically into the surface of a specimen. After the test force has been removed, the Brinell hardness is calculated from the remaining indentation diameter d :

$$HBW = 0.102 \frac{2F}{\pi D^2 (1 - \sqrt{1 - d^2/D^2})},$$

with load F in N, ball diameter D in mm, and mean indentation diameter d in mm.

Test loads range from 9.81 N to 29,420 N. Results obtained using balls of different diameters are only conditionally comparable, and any comparisons should be based on testing at identical force levels. Testing should always be performed using the largest possible ball, while load factors should be selected to obtain an indentation diameter of between $0.24 D$ and $0.6 D$. Table 1 lists the recommended load factors and ball diameters for a variety of materials as laid down in DIN EN ISO 6506-1 [3].

Table 1: Application of the Brinell hardness test

Material	Brinell hardness	Load factor $0.102 F/D^2$
Steel; nickel and titanium alloys		30
Cast iron (nominal diameter of ball must be 2.5, 5 or 10 mm)	< 140	10
	≥ 140	30
Copper and copper alloys	< 35	5
	35 to 200	10
	> 200	30
Light metals and their alloys	< 35	2.5
	35 to 80	5
		10
		15
	> 80	10
		15
Lead and tin		1
Sintered metals	see DIN EN ISO 4498 [4]	

In the Brinell-hardness designation, the numeric value is accompanied by the abbreviation for the method, the ball diameter in mm and the test force in N multiplied by a factor of 0.102 (e.g. 600 HBW 1/30).

High test loads producing deformations extending over a relatively wide surface area can be employed to gather data on materials with inconsistent structures. An advantage of the Brinell method is the relatively high degree of correlation between the Brinell hardness factor and the steel's tensile strength.

The required preparations and ensuing test procedures are more complex than those used for Rockwell testing.

Vickers hardness (DIN EN ISO 6507, [5])

This test method can be used for all metallic materials, regardless of hardness. It is especially suitable for testing minute and thin specimens, while the potential application range extends to include surface and case-hardened parts, as well as nitrided workpieces and parts carburized in nitrogen-based atmospheres.

The test tool is a square-based diamond pyramid with an apical angle of 136° . This is applied to the surface of the test specimen at an individually specified force F . The diagonals d_1 and d_2 of the indentation remaining on the test surface after test force F has been removed are measured. The Vickers hardness is obtained from the arithmetic mean value d of the two diagonals:

$$HV = \frac{2F \sin \frac{136}{2}}{d^2} \approx 0.1891 \frac{F}{d^2},$$

with test force F in N and arithmetic mean value d of diagonal lengths d_1 and d_2 in mm.

In the formula for Vickers hardness data, the actual test figure is accompanied by the abbreviation HV, the force in N (multiplied by a factor of 0.102) and, following a slash, the force application period (if other than the standard 15 seconds) in seconds (e.g.: 750 HV 10/25).

The test specimen's surface should be smooth and flat. DIN EN ISO 6507 [5] stipulates that correction factors be used to compensate for any error stemming from surface curvature. Test-load levels are selected with reference to either the thickness of the workpiece to be tested or the layer to be tested.

A major advantage of this test method is that there are virtually no limitations on using it to assess thin parts or layers. It allows the use of extremely small force levels to determine the hardness of individual structural sections. The Brinell and Vickers numbers correlate well up to approximately 350 HV. However, a certain minimal degree of surface consistency is necessary to ensure accurate results.

Knoop hardness (DIN EN ISO 4545, [6])

This process closely resembles the Vickers method. The equal-sided diamond tip in the Vickers test has a rhombic shape in the Knoop test. The test tool is designed to leave an impression in the form of a thin, elongated rhombus. The long diagonal d of the indentation remaining on the test surface after test force F has been removed is measured. The hardness value HK is calculated as follows:

$$HK = 1.451 \frac{F}{d^2},$$

with test force F in N and length d of the long diagonal in mm.

In the Knoop-hardness designation, the numeric value is accompanied by the abbreviation HK for the method, the test force in N multiplied by a factor of 0.102 and – if necessary separated by a slash – the application period of the test force in s (e.g. 640 HK 0.1/20).

The penetration depth is roughly 1/3 less than in the Vickers method, allowing evaluation of surface hardness in thin parts and layers. However, the test method also places heightened demands on the surface: The test must be performed on a polished, smooth and flat surface.

Knoop testing is often used on brittle materials such as, for example, ceramic or sintered materials.

Martens hardness (DIN EN ISO 14577, [7])

In this method, the penetration of a pyramid-shaped test tool into materials is recorded, in the course of which both the force and the travel during plastic and elastic deformation are measured. Martens hardness is defined as the ratio of test force F to surface A_s of the indenter calculated from penetration depth h , and given in N/mm^2 .

Martens hardness is denoted by the abbreviation HM, to which is appended the test force in N, the application time of the test force in s and the holding time of the test force in s (e.g. HM 0.5/20/20 $\triangleq$ 8,700 N/mm^2).

Scleroscope hardness

This dynamic measurement method is specially designed for heavy and large metal pieces. This process is based on the measurement of the rebound height of a steel indenter (hammer) featuring a diamond or hard metal tip. This is dropped from a stipulated height onto the surface of the test specimen. The rebound serves as the basis for determining the hardness.

The method is not standardized and there is no direct correlative relationship with any other hardness-testing method.

Heat-treatment processes

Heat treatment is employed to adapt the technological material properties of metallic components and tools to the relevant requirements. Such requirements can be processing properties to suit manufacturing and usage properties to suit function.

According to DIN EN 10052 [8], heat treatment involves "subjecting a workpiece in whole or in part to time and temperature cycles in order to bring about a change in its properties of its structure. If necessary, the chemical composition of the material can be changed during the treatment".

The process modifies the microstructure to achieve the hardness, strength, ductility, wear resistance, etc. required to withstand the stresses associated with static and dynamic loads. The most significant industrial processes are summarized in Table 2 (see DIN EN 10052, [8] for terminology).

Hardening

Hardening procedures produce a martensitic microstructure of extreme hardness and strength in ferrous materials such as steel and cast iron. These procedures consist of the individual stages known as austenitizing and cooling or quenching.

Through hardening

The workpiece being treated is heated to the austenitizing, or hardening temperature (Table 3), at which it is maintained until an austenitic structure emerges, and until an adequate quantity of carbon (released in the decay of carbides, such as graphite in cast iron) is dissolved in the treated material. After austenitizing, the workpiece is cooled or quenched at a rate sufficient for hardening, which can also occur in temperature stages. A complete conversion to the martensitic microstructure should be completed as quickly as possible. The necessary cooling process is determined by the chemical composition, the austenitizing conditions, the shape and dimensions, and the desired microstructure. Reference values for the required cooling rate can be found in the time-temperature transformation chart for the steel in question.

Table 2: Overview of heat-treatment processes

Hardening	Austempering	Draw tempering	Thermo-chemical treatment	Annealing	Precipitation hardening
Through hardening cross section	Isothermic transformation in the bainite stage	Tempering of hardened parts	Carburizing	Stress-free annealing	Solution treatment and aging
Surface hardening		Tempering above 540 °C for quenching and drawing	Carbonitriding	Recrystallization annealing	
Hardening of carburized parts (case-hardening)			Nitriding	Soft annealing	
			Nitro-carburizing	Spheroidization	
			Boron treatment	Normalizing	
			Chromating	Homogenizing	

The austenitizing temperature varies according to the composition of the material in question (for specific data, consult the DIN Technical Requirements for Steels). Table 3 above provides reference data. See DIN 17022 [9] Part 1 and Part 2 for practical information on hardening procedures for tools and components.

Not all types of steel and cast iron are suitable for hardening. The following equation describes the hardening potential for alloyed and unalloyed steels with mass carbon contents of between 0.15 % and 0.60 %, and can be applied to estimate the hardness levels achievable with a completely martensitic microstructure:

Max. hardness=(35+50x±2) HRC, (eq. 1)

where the carbon content in % by mass is to be applied for *x* here. If the microstructure does not consist entirely of martensite, then the maximum hardness will not be achieved.

When the carbon content exceeds 0.6 % by mass, it may be assumed that the material's structure contains untransformed austenite (residual austenite) in addition to the martensite. Greater amounts of residual austenite can have a deleterious effect on achievable hardness and reduce wear resistance. In addition, residual austenite is metastable, i.e. there exists a potential for subsequent transformation to martensite at temperatures below room temperature or under stress. This can give rise to unwanted changes in the workpiece's dimensions and shape. Prompt low-temperature follow-up procedures or draw-tempering operations at over 230 °C can be useful in cases where residual austenite is an unavoidable product of the hardening procedure.

Temperature differences occur between the edge and core of the workpieces during cooling and quenching. With greater cross-sections, the reduced cooling rate in the core can give rise to a decrease in hardness as the distance to the surface increases. There is a hardness progression or gradient. The hardness

Table 3: Standard austenitizing temperatures

Type of steel	Quality specification [10, 11, 12]	Austenitizing temperature in °C
Unalloyed and low-alloy steels	DIN EN 10083-1 10083-2 10083-3	780...950 750...820
< 0.8 % by mass of C ≥ 0.8 % by mass of C	10085 –	
Cold- and hot-working tool steels	DIN EN ISO 4957	780 to 1150
High-speed tool steels		1150 to 1300

gradient is obtained from the material composition and the hardenability dependent on the austenitizing conditions (testing described in DIN EN ISO 642, [13]). In this case, it is necessary with regard to the required hardness to use material with sufficient hardenability. Information on choice of steel based on hardenability can be found in DIN 17021 [14].

DIN EN ISO 18265 [15] defines the method for using hardness as the basis for estimating tensile strength R_m . This method can only be applied in cases whether the surface and core hardnesses are virtually identical.

During hardening, the transformation of the microstructure to the martensitic state is combined with an increase in volume. With regard to the initial state, the volume increases by approximately 1%. This equates to a change in length of approximately 0.3%.

The changes in volume associated with rearrangements of the structure and thermal gradients during cooling give rise to stresses, which in turn result in distortion in the form of changes in dimension and shape. The stresses that remain in the workpiece after hardening are called internal stresses. The edge of a hardened workpiece tends to be subject to internal tensile stress, while the core tends to be subject to internal compressive stress.

Table 4: Comparison of power densities when heating with different sources

Energy source	Normal power density in W/cm ²
Laser beam	10 ³ to 10 ⁴
Electron beam	10 ³ to 10 ⁴
Induction (with medium- or high-frequency alternating current, or high-frequency pulses)	10 ³ to 10 ⁴
Flame heating	1 · 10 ³ to 6 · 10 ³
Plasma beam	10 ⁴
Molten saline solution (convection)	20
Air, gas (convection)	0.5

Surface hardening

This process is especially suited for integration within large-scale manufacturing operations, and can be adapted to fit the rhythm of the production line.

Heating and hardening are restricted to the surface, thereby minimizing alterations in shape and dimensions. Heating is generally provided by high- or medium-frequency alternating current (induction hardening) or by a gas burner (flame hardening). Friction (friction hardening) and high-energy beams (e.g. electron or laser beams) can also provide the heat required for austenitizing. Table 4 provides an overview of the specific heat energies for the individual procedures.

These methods can be used to treat both linear and flat surfaces, meaning that the parts can be heated either while stationary or in motion. The heat source itself can also be moved. Rotation is the best way of dealing with radially symmetrical parts, as it ensures concentric hardening. Either immersion or spraying arrangements can be applied for quenching. Information on performing surface hardening can be found in DIN 17022-5 [9].

Heat rise is rapid, so the temperatures must be 50 to 100 °C higher than those used in furnace heating so as to compensate for the shorter dwell period. The procedure is generally employed with low-alloy or unalloyed steels with mass carbon contents of 0.35 to 0.60 %. However, surface hardening processes can also be applied with alloyed steels, cast iron and rolling-bearing steels. The parts can be heat-treated to provide a combination of improved base strength and high surface hardness, making them suitable for high-stress applications (recessed edges, bearing surfaces, cross-sectional transitions).

Surface hardening generally results in internal compression stresses along the edge. This leads to increased fatigue resistance, especially when notched parts are exposed to inconstant vibration stress. The stress in Figure 2 corresponds to bending stress. The higher stressability results from the fact that the stress state (resulting stress) is reduced from the superposition of bending stress and internal stress.

The relationship defined in equation 1 can be employed to estimate the potential surface hardness. There is a substantial reduction in hardness between the surface and the unhardened core region. The surface hardness depth *SHD* – the depth at which 80 % of the minimum Vickers surface hardness is found – can be derived from the hardness progression curve (see DIN EN 10328, [16]).

Austempering

The object of this process is to achieve a bainite microstructure. This microstructure is not as hard as martensite, but does display greater ductility, as well as smaller changes in specific volume.

After austenitizing (see hardening), the parts for austempering are first cooled to a temperature of 200 to 350 °C (depending upon the exact composition of the material) at the required rate. The parts are then held at this temperature until the microstructure's transformation into bainite has been completed. Cooling to the transformation temperature is usually performed in molten saline solutions (a typical salt is a mixture of potassium nitrate and sodium nitrite). Following the microstructure transformation the parts can be cooled to room temperature (no special procedure required).

Austempering is an excellent alternative for parts whose geometrical configuration makes them sensitive to distortion or cracks, or in which high ductility is required together with substantial hardness, or which should combine hardness with a low level of residual austenite.

Application

Cylinder heads in modern diesel high-pressure pumps for common-rail systems which must be able to withstand high wear and internal compressive stresses are austempered.

Draw tempering

Draw tempering of hardened components and tools is employed to increase their deformability and to reduce their risk of cracking. According to DIN EN 10052 [8], draw tempering involves heating the part in question once or repeatedly to tempering temperature, holding the part at this temperature, and then allowing it to cool appropriately. Draw tempering is performed between room temperature and A_{c1} temperature, i.e. the temperature at which austenitic structural constituents would be created.

Tempering at temperatures as low as 180 °C is enough to reduce the hardness of unalloyed and low-alloy steels by approx. 1 to 5 HRC. The individual materials respond to higher temperatures with specific characteristic hardness loss. Figure 3 shows a characteristic tempering curve for typical types of steel. This graph illustrates the fact that the hardness of steels which are alloyed with special carbide-forming elements (Mo, V, W) – such as, for example, hot-working or high-speed tool steel – is increased by tempering in a temperature range between 400 and 600 °C to values which can be above the quenching hardness (secondary hardening).

The mutual relationships between tempering temperature on the one side, and hardness, strength, yield point, fracture contraction and elongation at fracture on the other, can be taken from the tempering diagrams for the various steels (see e.g. DIN EN 17021, [17]).

Generally speaking, draw tempering reduces hardness and strength and increases deformability. Internal stresses

Figure 2: Cyclically alternating stress according to surface-layer hardening

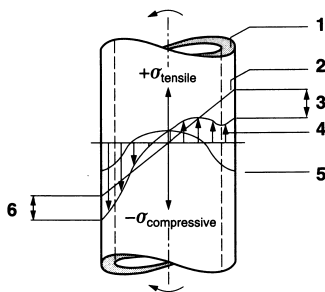
+ σ Tensile, – σ Compressive.

1 Case layer, 2 Bending stress,

3 Reduction of tensile stress,

4 Resulting stress, 5 Internal stress,

6 Increase in compressive stress.



can also be reduced at tempering temperatures in excess of 300 °C.

The specific volume decreases when structures which are free of residual austenite are tempered. In the case of structures containing residual austenite, however, an increase in volume occurs during the transformation from residual austenite to martensite. Hardness increases, deformability is reduced, and new internal stresses can be created. The risk of cracking also increases.

It must be remembered that steels alloyed with manganese, chromium and nickel, or combinations of these elements, should not be tempered at temperatures of 350 to 550 °C, as brittleness could result. When these types of materials are cooled from tempering temperatures above 550 °C, the transition through this critical range should also be effected as rapidly as possible (see DIN 17022 Parts 1 and 2, [9] for additional information). This tempering sensitivity can be avoided by adding molybdenum or tungsten by alloying.

Quenching and drawing

Quenching and drawing involves a combination of hardening and tempering at a temperature which is generally between

540 °C and 680 °C. This procedure is designed to achieve an optimal relationship between strength and ductility. It is applied in cases where extreme ductility or malleability is required.

Particular care must be devoted to avoiding brittleness in the quench and draw operation (see above).

Annealing

Annealing can be applied to optimize certain operational or processing characteristics of parts. With this method, the parts are heated to a specific temperature and maintained there for an adequate period before being cooled to room temperature. The most important annealing processes in terms of technology are described in the following.

Stress-free annealing

Depending on the precise composition of the parts, stress-free annealing is carried out at temperatures ranging from 450 °C to 650 °C. The specific object is to achieve a reduction in internal stress in components, tools and castings.

Recrystallization annealing

Recrystallization annealing is applied with cold-form parts. The goal is to restructure the grain pattern in order to prevent increased hardening, thereby facilitating subsequent machining work.

The temperature requirement depends upon the composition of the material and the degree of deformation: it lies between 550 °C and 730 °C for steel.

Soft annealing and spheroidization

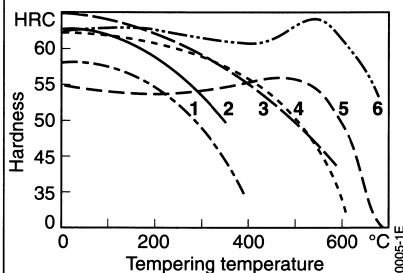
The purpose of soft annealing is to improve the machinability of material states which are difficult to cut or difficult to cold-form. The process involves heating the material in question to temperatures in excess of 600 °C, as briefly as possible above the A_{c1} temperature for the relevant steel, holding at this temperature, and slowly cooling to room temperature.

The temperature requirement is determined by the material's composition. It ranges from 650 °C to 850 °C for steel, and lower for nonferrous materials.

Spheroidization of cementite is applied when a microstructure with a granular

Figure 3: Tempering response of various types of steel

- 1 Unalloyed tempering steel (C45),
- 2 Unalloyed cold-working tool steel (C80W2),
- 3 Low-alloy cold-working tool steel (105WCr6),
- 4 Alloyed cold-working tool steel (X165CrV12),
- 5 Hot-working tool steel (X40CrMoV51),
- 6 High-speed tool steel (S6-5-2).



carbide pattern is desired. The cementite loses strength as a result of the annealing temperature and can pursue its striving for a body with as small a surface as possible (of the ball). If the initial structure is martensite or bainite, the result will be an especially homogeneous carbide distribution.

Normalizing

Normalizing is carried out by heating the parts to austenitizing temperature and then allowing them to gradually cool to room temperature. In low-alloy and unalloyed steels, the result is a structure consisting of ferrite and perlite. This process is essentially employed to reduce grain size, reduce the formation of coarse grain patterns in parts with limited reshaping, and to provide maximum homogeneity in the distribution of ferrite and perlite.

Precipitation hardening

This process combines solution treatment with aging at ambient temperature. The parts are heated and then maintained at a temperature to bring precipitated structural constituents into a solid solution, and quenched at room temperature to form a supersaturated solution. The aging process comprises one or several cycles in which the material is heated and held at above-ambient temperatures ("hot aging"). In this process, one or several phases, i.e. metallic bonds between certain base alloys, are formed and precipitated in the matrix.

The precipitated particles enhance the hardness and strength of the base microstructure. The actual characteristics are determined by the temperature and duration of the aging process (option of mutual substitution). Exceeding a certain maximum will usually reduce the strength and hardness of the final product.

Precipitation hardening is mostly applied for nonferrous alloys, but some hardenable steels (maraging steels) can also be processed.

Application

Steels which can be precipitation-hardened are used for example in rail-pressure sensors in common-rail systems.

Thermochemical treatment

In thermochemical treatment, the exchange of substances with suitable agents effects a change in the chemical composition of the base material. Specific function properties can be adapted by the diffusion of specific elements into the surface layer. Of particular importance for this process are the elements carbon, nitrogen and boron.

Carburizing, carbonitriding and case hardening

Carburizing increases the carbon content in the surface layer, while carbonitriding supplements the carbon enrichment with nitrogen. This process usually takes place at temperatures ranging from 750 to 1,050 °C in gases which give off carbon or nitrogen as a result of their disintegration by heat or excited in the plasma. The actual hardening is performed subsequently, either by quenching directly from the carburizing or carbonitriding temperature (direct hardening), or by allowing the parts to cool to room temperature (single hardening), or by allowing them to cool to a suitable intermediate temperature (e.g. 620 °C) prior to reheating (hardening after isothermic conversion). This process produces a martensitic surface layer, while the degree of martensite at the core is a function of hardening temperature, hardness and part thickness.

Specific temperatures can be selected for either surface hardening in the upper layers with higher carbon content (case refining), or for the non-carburized core (core refining) (see DIN 17022 Part 3, [9]). Carburizing and carbonitriding produce a characteristic carbon declivity, with levels dropping as the distance from the surface increases (carbon curve). The distance between the surface and the point at which the mass carbon content is still 0.35 % is normally defined as the carburization depth.

The length of the carburizing or carbonitriding process depends upon the required carburization depth, the temperature and the atmosphere's carbon-diffusion properties.

Typical surface carbon contents during case hardening range between 0.5 and 0.85 % carbon by mass in order to achieve

sufficient surface hardness. The carbon concentration essentially determines the surface hardness.

Excessive concentrations of carbon can lead to residual austenite or carbide diffusion, which could have negative effects on the performance of case-hardened parts in actual use. Control of the atmosphere's carbon level is thus extremely important in process management.

The most common carburizing processes used today are gas carburizing and vacuum carburizing. In the case of gas carburizing, the carbon level of the furnace atmosphere is regulated in such a way that the surface layer of the workpiece assumes the desired carbon content. A state of equilibrium is established with the surrounding furnace atmosphere. In the case of vacuum carburizing, on the other hand, the carbon content cannot be regulated. Here, the defined surface carbon content is adjusted by means of multistage carburizing. In a first stage, carburizing is performed to a very high surface carbon content in the range of material saturation. In the subsequent stage, this high carbon content is reduced by diffusion to the desired level. In practice, vacuum-carburizing processes consist of several consecutive carburizing and diffusion stages.

The gradient defining the relationship between hardness and depth corresponds to the carbon concentration curve. The case-hardening depth *CHD* can be taken from this. DIN EN ISO 2639 [18] defines this as the vertical distance to the surface to the layer which has a Vickers hardness of 550 HV 1.

The case-hardened part generally exhibits compression tension at the surface, and tensile stresses at the core. As with surface-hardened materials, this distribution pattern provides enhanced resistance to vibration loads.

In carbonitriding, nitrogen is also absorbed; it serves to improve the material's tempering properties, increase its durability and enhance its wear resistance. The positive effects are especially pronounced with unalloyed steels.

For additional, more detailed information on case hardening procedures, consult DIN 17022 Part 3 [9], and Information

Sheet 452 of the Steel Information Center, Düsseldorf [19].

Application

Injection nozzles subject to wear for Bosch common-rail systems which must be able to withstand high wear and internal compressive stresses are case-hardened by vacuum carburizing.

Nitriding and nitrocarburizing

Nitriding is a thermal treatment process (temperature range: 400 to 600 °C) which can be used to enrich the surface layer of virtually any ferrous material with nitrogen. In nitrocarburizing, a certain amount of carbon is diffused into the material to form nitrogen at the same time.

Molecular nitrogen, as is present in this temperature range in gaseous nitrogen, cannot diffuse into metallic materials. It is therefore necessary to offer diffusible nitrogen via suitable donor media. In practical applications, nitriding and nitrocarburizing processes are performed in gas atmospheres containing ammonia, in plasma containing nitrogen, or even in molten saline solutions containing cyanate. While ammonia gas releases diffusible nitrogen during its thermal disintegration, nitrogen is ionized in the plasma in order to split the molecules and facilitate the diffusion of nitrogen atoms.

The enriching nitrogen in the surface layer causes the precipitation of nitrides, in response to which the surface layer hardens. In the final analysis, this results in greater resistance to wear and corrosion, and in greater endurance strength.

As the process employs relatively low temperatures, there are no volumetric changes of the kind associated with transformations in the microstructure, so that changes in dimensions and shape are minute.

The nitrided region consists of an outer layer, several millimeters in depth, and a transitional white layer, the hardness of which may be anywhere from 700 HV to over 1,200 HV, depending on the composition of the material. Still deeper is a softer diffusion coating extending several tenths of a millimeter in which the nitrogen content decreases as the distance from the surface increases. The thickness

of the individual layers is determined by the temperature and duration of the treatment process. The process produces a hardness gradient (similar to that which results from surface and case hardening); this gradient furnishes the basis for determining the nitriding hardness depth *NHD*. DIN 50190 Part 3 [20] defines this as the vertical distance from the surface to the point where the hardness corresponds to a defined limit. This limit hardness is usually actual core hardness + 50 HV 0.5.

The material's resistance to wear is essentially determined by the white layer, which contains up to 10 mass components of nitrogen in %. The nitriding hardness depth and the surface hardness determine the material's resistance to alternating cyclic stress (for additional details, see DIN 17022 Part 4, [9], and Information Sheet 447 from the Steel Information Center, Düsseldorf, [21]).

The corrosion resistance of nitrided or nitrocarburized workpieces can be significantly increased by postoxidation in water vapor or other suitable gases, or in molten saline solutions at temperatures ranging between 350 °C and 550 °C.

Application

Injection nozzles which are used, for example, in Bosch common-rail systems and which must be able to withstand extreme operating temperatures are gas-nitrided.

Components for windshield and rear-window wiper systems are nitrocarburized and postoxidized to increase their resistance to corrosion and wear. This is how these components get their typical black color.

Boron treatment

Boron treatment involves enriching the surface layer of ferrous materials with boron. Depending upon duration and temperature (normally 850 to 1,000 °C) of the treatment, a white layer of 30 µm to 0.2 mm in depth with a hardness of 2,000 to 2,500 HV is produced. It consists of iron-boron.

Boron treatment is particularly effective as a means of protecting against abrasive wear. However, the comparatively high process temperature leads to relatively

large changes in shape and dimensions, meaning that this treatment is only suitable for applications in which large tolerances can be accepted.

Application

Partly boron-treated tool holders with high wear resistance are used in Bosch hammer drills.

References

- [1] DIN EN ISO 18265: Metallic materials – Conversion of hardness values.
- [2] DIN EN ISO 6508: Metallic materials – Rockwell hardness test.
- [3] DIN EN ISO 6506: Metallic materials – Brinell hardness test.
- [4] DIN EN ISO 4498: Sintered metal materials, excluding hardmetals – Determination of apparent hardness and microhardness.
- [5] DIN EN ISO 6507: Metallic materials – Vickers hardness test.
- [6] DIN EN ISO 4545: Metallic materials – Knoop hardness test.
- [7] DIN EN ISO 14577: Metallic materials – Instrumented indentation test for hardness and materials parameters.
- [8] DIN EN 10052: Vocabulary of heat treatment terms for ferrous products.
- [9] DIN 17022: Heat treatment of ferrous materials – Methods of heat treatment. Part 1: Hardening, austempering, annealing, quenching, tempering of components. Part 2: Hardening and tempering of tools. Part 3: Case hardening. Part 4: Nitriding and nitrocarburizing. Part 5: Surface hardening.

[10] DIN EN 10083: Steels for quenching and tempering. Part 1: General technical delivery conditions.

Part 2: Technical delivery conditions for non-alloy steels.

Part 3: Technical delivery conditions for alloy steels.

[11] DIN EN 10085: Nitriding steels – Technical delivery conditions.

[12] DIN EN ISO 4957: Tool steels.

[13] DIN EN ISO 642: Steel – Hardenability test by end quenching (Jominy test).

[14] DIN 17021: Heat treatment of ferrous materials.

[15] DIN EN ISO 18265: Metallic materials – Conversion of hardness values.

[16] DIN EN 10328: Iron and steel – Determination of the conventional depth of hardening after surface hardening.

[17] DIN 17021-1: Heat treatment of ferrous materials; material selection, steel selection according to hardenability.

[18] DIN EN ISO 2639: Steels – Determination and verification of the depth of carburized and hardened cases.

[19] Information Sheet 452 of the Steel Information Center, Düsseldorf: "Einsatzhärten", Edition 2008.

[20] DIN 50190-3: Hardness depth of heat-treated parts; determination of the effective depth after nitriding.

[21] Information Sheet 447 of the Steel Information Center, Düsseldorf: "Wärmebehandlung von Stahl – Nitrieren und Nitrocarburieren", Edition 2005.

Corrosion and corrosion protection

Corrosion processes

Corrosion is the damage to a metal as a result of a reaction with substances in the environment. Corrosion processes always contain interphase reactions. An example of this type of reaction is metal scaling, i.e. oxidation in hot gases. As it proceeds, metal atoms oxidize to form nonmetallic compounds in a process emanating from the affected material's surface. In thermodynamic terms, the process can be viewed as an entropic transition from an ordered, high-energy state into a less-ordered state of lower energy, and consequently greater stability. The process takes place of its own accord.

The following deals exclusively with the corrosion that occurs at the phase boundary between the metal and aqueous phases (electrolyte), generally referred to as electrochemical corrosion.

Corrosive attack

Anodic subprocess

Two essentially distinct reactions occur in electrochemical corrosion: in the anodic subprocess, the directly visible corrosion process, the metal passes into the oxidized state as described in the reaction equation

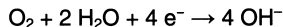


freeing an equivalent number of electrons (Figure 1). The metal ions thus formed can

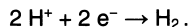
either be dissolved in the electrolyte, or can precipitate out on the metal as corrosion products (e.g. rust) after reacting with constituents in the attacking medium [1].

Cathodic counter-reactions

This anodic subprocess can continue only as long as the electrons it produces are consumed in a second process. This second subprocess is a cathodic subreaction. In neutral or alkaline media oxygen is reduced in accordance with



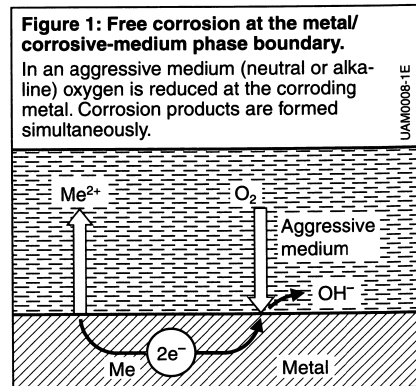
to hydroxyl ions, which in turn are able to react with the metal ions. In acidic media the hydrogen ions are reduced via the formation of free hydrogen, which escapes as a gas according to the following formula:



Electrochemical series of metals

Metals are often ranked in "electrochemical series of metals" corresponding to consecutively higher standard potentials (see also Electrochemistry). As the value of standard potential increases the metals are referred to as "more noble", and as the value decreases as "more base".

It should be emphasized that the table ("Standard electric potentials of metals") is limited to thermodynamic values, and does not reflect the effects of corrosion kinetics, for example as encountered in the formation of protective layers. Most corrosion-resistant structural materials, however, develop protective layers, which are crucial to corrosion resistance. Thus, for example the materials aluminum, zinc and titanium seem to be very base metals, but are highly resistant due to the formation of protective layers. The formation of protective layers is exploited in the field of electrochemical corrosion protection.



Types of corrosion

General surface corrosion

General surface corrosion is a type of corrosion with uniform attrition of material over the entire boundary surface between the material and the attacking medium. This is a very frequent type of corrosion in which the material penetration rate (removal depth) can be calculated per unit of time based on the corrosion current.

Pitting corrosion

Pitting corrosion is a limited localized attack by a corrosive medium which penetrates the material by forming holes, or pits, whose depth is almost always greater than their diameter. Practically no material is removed from the surface outside the pitted areas. Pitting corrosion is frequently caused by chloride ions (e.g. from table salt).

Contact corrosion

When two different metals moistened by the same medium are in mutual electrical contact, a cathodic subprocess occurs at the more noble metal, while the anodic subprocess progresses at the baser material. This is called contact corrosion.

Crevice corrosion

Crevice corrosion is a corrosive attack primarily occurring in narrow crevices, caused by concentration differences in the corrosive medium, e.g. as a result of long oxygen diffusion paths. This type of corrosion generates potential differences between the crevice extremities, leading to intensified corrosion in more poorly ventilated areas.

Stress corrosion cracking

This is corrosion stemming from the simultaneous concerted action of a corrosive medium and mechanical tensile stress (which can also be present as internal stress in the object itself). Intergranular or transgranular fissures form, in many cases without the appearance of visible corrosion products.

Vibration corrosion cracking

Vibration corrosion cracking is corrosion caused by the simultaneous effects of a corrosive medium and mechanical fatigue stress, e.g. caused by vibrations. Transgranular cracks are formed, frequently without visible deformations.

Intergranular and transgranular corrosion

This involves types of corrosion characterized by selective formation along the grain boundaries or roughly parallel to the deformation plane in the grain interior.

Dezincification

Selective dissolution of zinc from brass, leaving behind a porous copper structure.

Corrosion testing

Electrochemical testing procedures

The primary tool for electrochemical corrosion testing is the potentiostat. Here metals or other electrically conductive materials are analyzed in different corrosive media. The typical setup with three electrodes is shown in Figure 2. This arrangement is essentially used to determine – as well as the corrosion potentials of the corroding materials – primarily the corrosion currents. In the case of uniform surface corrosion, the corrosion-current parameters are then used for the definition of the attrition mass and removal depth per unit of time.

The electrochemical testing procedures are a valuable supplement to the non-electrochemical methods, because they provide an understanding of corrosion mechanisms that occur. In addition to the small amount of corrosive medium required, another advantage of electrochemical procedures over non-electrochemical methods is that they provide quantitative data on attrition rates.

The potentiostat is used to alter the potential of the material to be analyzed and measure the flowing current. Evaluation enables reaction mechanisms and above all corrosion current densities to be determined.

Polarization-resistance measurement

In the case of free corrosion the corrosion current is determined from the polarization resistance (slope of the cumulative current-density/potential curve); testing entails subjecting the metal to minimal, alternating anodic and cathodic pulses.

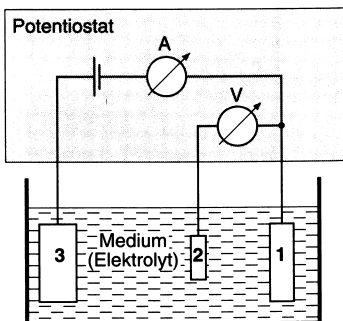
Impedance spectroscopy

Electrochemical Impedance Spectroscopy (EIS) is employed to examine corrosion mechanisms. This alternating-current technique determines the AC resistance (impedance) of an electrochemical test object as a function of frequency. A low-amplitude sinusoidal alternating voltage is superimposed on the working electrode's potential, and the current response is measured. To interpret the measurement the system is approximated in the form of an equivalent network. By way of example, Figure 3 shows the equivalent network for the system, consisting of metal, coating, and medium.

Impedance elements (resistances, capacitances, inductances) are assigned to the sequences (e.g. electron transfer) at the phase boundary. In this simple example from Figure 3 the coating is – along similar lines to the phase boundary of material to corrosive medium – described as capacitor and resistor.

Figure 2: Schematic setup for measuring the current density/potential curves with a potentiostat

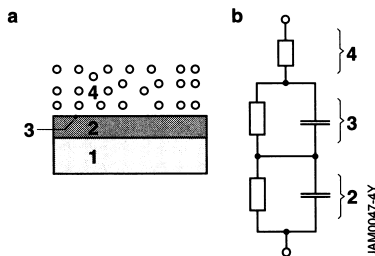
- 1 Working electrode,
- 2 Reference electrode,
- 3 Counter-electrode.



UAM0149-1D

Figure 3: Evaluation of EIS data (electrochemical impedance spectroscopy)

- a) Arrangement,
 - b) Equivalent network.
- 1 Metal,
 - 2 Coating,
 - 3 Phase boundary metal to medium,
 - 4 Corrosive medium.



UAM0047-4Y

Direct conclusions can then be drawn from the equivalent networks and the impedance elements about various characteristics, such as the effectiveness of corrosion-protection measures, porosity, thickness, a coating's water-absorption ability, the effectiveness of inhibitors, the corrosion rate of the base metal, and so on.

Contact-corrosion current measurement

When contact corrosion is measured the two affected metals are immersed in the same corrosive medium. During measurement they are not directly contacted, but instead connected via the potentiostat. Here the contact-corrosion potential and the contact-corrosion current flowing between the two metals are directly measured and the curve plotted against the measurement period.

Non-electrochemical corrosion-testing procedures

In non-electrochemical testing procedures test specimens essentially are first exposed to a corrosive environment (exposure tests) and then the change in the test specimens is assessed.

Simple standard tests (e.g. corrosion load by neutral salt spray fog testing) are conducted with the aim of testing quality or comparing corrosion-protection measures and materials. The emphasis is placed on using tests that can uncover weak points in corrosion-protection measures and describe different qualities.

After corrosion load is completed, it is possible by comparing images in accordance with DIN EN ISO 4628-3 [2] to assign rust levels (degrees of rusting) which are defined according to rust coverage or surface perforation (Table 1).

Table 1: Rust level and rust surface acc. to DIN EN ISO 4628-3

Rust level	Rust surface in %
R _i 0	0
R _i 1	0.05
R _i 2	0.5
R _i 3	1
R _i 4	8
R _i 5	40...50

For the purpose of service-life assessment, testing procedures have evolved to reflect specific requirements, e.g. motor-vehicle testing. These are as a rule corrosion loads with cyclically alternating loading by salt spray fog, dry phase and moist phase. These tests provide reliable indices of projected service life under normal operating conditions by using short-term exposure in extremely harsh conditions to simulate long-term stresses in the real world (e.g. testing in accordance with Bosch standard N42AP 226 "Climatic tests – Tightened life – Corrosion testing").

For a further adaptation of corrosion load to practical operation products are operated under tightened conditions (e.g. vehicle tests of the different automobile manufacturers).

In addition to the corrosion loads for life testing the focus is on assessing the function of the products.

Corrosion protection

The manifestations and mechanisms of corrosion are many and varied, so widely differing methods can be adopted to protect metals against corrosion attack. Corrosion protection means intervening in the corrosive process with the object of reducing the rate of corrosion in order to prolong the service life of the components.

Corrosion protection can be achieved by applying four basic principles:

- Measures in planning and design: the choice of suitable materials and the suitable structural design of components.
- Measures that intervene in the corrosive process by electrochemical means.
- Measures that separate the metal from the corrosive medium by protective layers or coatings.
- Measures that influence the corrosive medium, for example, the addition of inhibitors to the medium.

Corrosion protection by means of suitable structural design

Material selection

Selecting suitable materials which feature optimum resistance to corrosion under the expected conditions can be of consid-

erable assistance in avoiding corrosion damage. When the costs that would otherwise be incurred for upkeep and repair are factored into the long-term cost of ownership equation, a more expensive material can often be the more cost-effective alternative.

Design

Design measures, too, are of major importance. A great deal of skill and expertise goes into design, particularly regarding the connections between parts that are made of the same material or different materials.

Corners and edges of sections are difficult to protect, and this is where corrosion can easily attack. A favorable installation position can avoid corrosion (Figure 4).

Beads and welts can trap dirt and moisture. Suitable surfaces and drain openings can help avoid this problem (Figure 5).

In the case of welds the structure is altered in a detrimental way. In order to avoid crevice corrosion, welds have to be smooth and free of gaps (Figure 6).

Contact corrosion can be avoided by joining same or similar metals, or by installing washers, spacers, or sleeves to ensure that both metals are electrically insulated (Figure 7).

Figure 4: Good and bad installation positions for profile sections

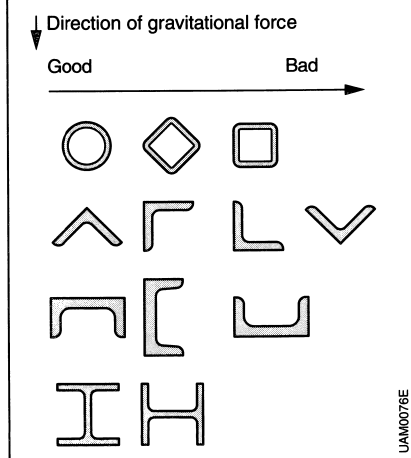
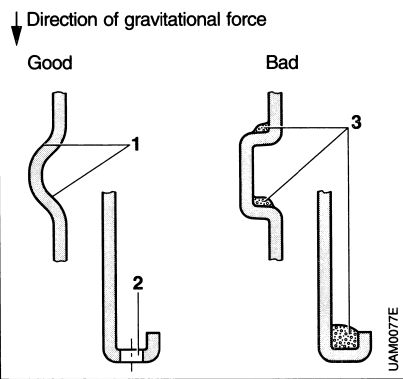


Figure 5: Design of beads and welts

- 1 Sloping faces (foreign matters slide off),
- 2 Wide gap with drain,
- 3 Deposits of dirt and moisture.



Electrochemical processes

The schematic current-density/potential curves for a metal suitable for passivation coating (e.g. by oxide layer on stainless steel) show how these processes work (Figure 8). The current-density values arranged in ascending order on the y-axis represent anodic currents corresponding to the corrosion reaction defined in the equation



In contrast, the descending current-density values represent cathodic currents.

As can be seen from the diagram, corrosion only occurs when potentials $E_a \leq E \leq E_p$ and $E \geq E_d$. Therefore corrosion can be suppressed by applying an external voltage. There are two basic ways of doing this.

Cathodic protection

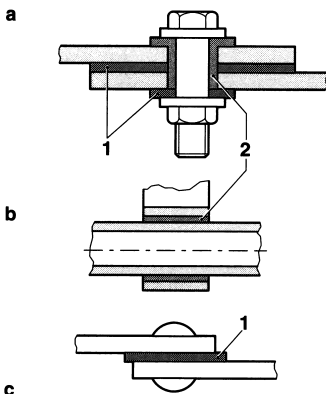
For cathodic protection, the potential is shifted so far toward the left that no anodic currents flow, leading to $E < E_a$. As an alternative to applying voltage from an external source, it is also possible to shift the potential by contacting a base metal to act as a "sacrificial" reactive anode.

Anodic protection

Another option is, by applying an external voltage, to shift the potential of the threatened electrode into the passive range, i.e. into the potential range between E_p and

Figure 7: Electric insulation to avoid contact corrosion

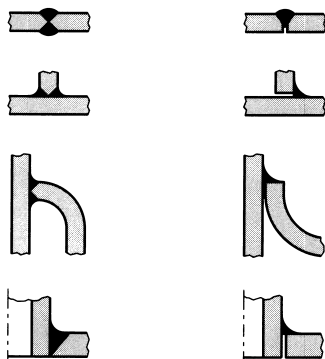
- a) Screwed/bolted joint,
 - b) Retainer,
 - c) Riveted joint.
- 1 Insulating washer,
2 Insulating spacer.



UAM0079Y

Figure 6: Design-related crevices in welds and how they can be avoided

Good (through fusion) Bad (crevice)

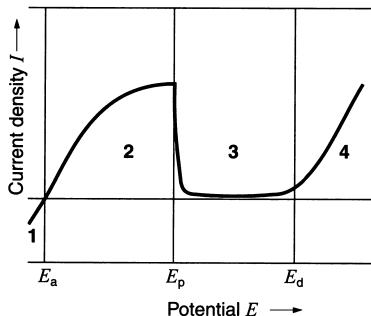


UAM0078E

Figure 8: Schematic curve of current density as a function of the potential for a metal which can be passivated

- 1 Cathodic protection, occurs when $E < E_a$,
- 2 Corrosion without protection,
- 3 Anodic protection, occurs when $E_p < E < E_d$,
- 4 Transpassive range.

E_a Free corrosion potential of the metal in an active state,
 E_p Passivating potential,
 E_d Breakthrough potential.



UAM0011-3E

E_d . This is called anodic protection. The anodic currents, which flow in the passive range, are less than those in the active range by exponential powers of between three and six, depending on the type of metal and the corrosive medium. The result is excellent protection for the metal.

However, the potential should not exceed E_d , as oxygen would be produced in this transpassive range, potentially leading to higher rates of oxidation. Both of these effects would cause the current to increase.

Coatings

Coatings inhibit corrosion by forming protective films which are applied to the surface of the metal to be protected, where they resist attack by the corrosive medium (see also Deposits and coatings).

Inhibitors

Inhibitors are substances added to the corrosive medium in low concentrations (up to a maximum of several hundred ppm) for absorption at the surface of the protected metal. They drastically reduce the rate of corrosion by blocking either the anodic or the cathodic subprocess (frequently blocking both subprocesses simultaneously). Organic amines and amides of organic acids are the most frequent inhibitors. In automotive applications, for example, inhibitors are used in fuel additives. They are also added to anti-freeze in order to inhibit corrosion damage in the coolant circuit.

Vapor-phase inhibitors

Vapor-phase inhibitors (VPI), or volatile corrosion inhibitors (VCI), are organic substances of moderate vapor pressure. They are frequently enclosed in special-purpose packing materials or as solvents in liquids or emulsions with oil. The inhibitors evaporate or sublime over the course of time, and are adsorbed as monomolecules on the metal, where they inhibit (i.e. slow down or stop) either anodic or cathodic corrosion subreactions, or both at once. Dicyclohexylamin nitrite is a typical example.

For optimal effectiveness, the inhibitor should form a sealed coating extending over the largest possible surface area.

Standard commercial vapor-phase inhibitors are generally a combination of numerous components capable of providing simultaneous protection for several metals or alloys; exceptions: cadmium, lead, tungsten and magnesium.

Vapor-phase inhibitors provide only temporary protection for metallic products during storage and shipping. They must be easy to apply and remove. The drawback of vapor-phase inhibitors is that they are potentially injurious to health.

References

- [1] DIN 50919:2016-02: Corrosion of metals – Corrosion investigations of bimetallic corrosion in electrolytic solutions.
 - [2] DIN EN ISO 4628-3:2016-07: Paints and varnishes – Evaluation of degradation of coatings – Designation of quantity and size of defects, and of intensity of uniform changes in appearance.
- Part 3: Assessment of degree of rusting (ISO 4628-3:2016); German version EN ISO 4628-3:2016.

Deposits and coatings

Deposits and coatings are used to adapt the surface properties of components to particular requirements ("surface engineering", Table 1). It is thus possible, for example, for a component to be made from a tough, low-cost material which still has a hard and wear-resistant surface. The main applications for deposits and coatings are

- Corrosion protection (as well as functional retention, often for decorative reasons)
- Wear protection
- Joining and bonding techniques (plug-in, welded, soldered, bonded, and crimped contacts).

In coating systems, distinctions are made between the following types of coating:

- Deposits in which one layer is applied
- Conversion coatings in which the functional coating is produced by chemical or electrochemical conversion of the base material, and
- Diffusion coatings in which the functional coating is produced by diffusion of atoms or ions into the base material.

Deposits

Electrodeposits

Electrodeposits are deposited using an external power source. The workpieces to be coated are immersed into an electrolyte (Figure 1) which contains ions of the metal to be deposited. The metal ions consumed during the deposition process are either supplied by the anode solution or added to the electrolyte (when inert anodes are used). The progression and distribution of the field lines influence the liner-thickness distribution. A uniform layer thickness can be achieved by means of an optimized configuration of anodes and screens.

Electrodeposits are widely used in corrosion protection, in wear protection and for electrical contacts. Some important deposits are described below.

Zinc and zinc alloys

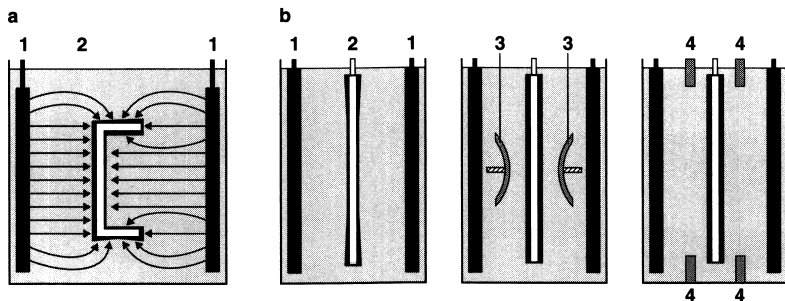
Electrolytically deposited zinc coatings are widely used as corrosion protection for steel components and are a cathodically effective means of corrosion protection (sacrificial effect). To increase the corrosion-protection effect, zinc deposits are passivated in a chrome(VI)-free solution. Zinc alloys, such as zinc-nickel with

Figure 1: Electrodeposits

a) Progression and distribution of electrical field lines between anode and cathode as caused by nonuniform liner-thickness distribution.

b) Comparison of liner-thickness distribution during electroplating without auxiliary agent, with auxiliary anodes and with screens (schematic representation).

1 Anode, 2 Cathode, 3 Auxiliary anode, 4 Screen.



approx. 15 % nickel offer a significantly higher degree of corrosion protection.

Nickel

Electrolytically deposited nickel coatings offer limited corrosion protection with an attractive visual appearance. One area of automotive engineering in which this is chiefly used is the nickel-plating of spark-plug shells.

Chrome

In the case of chrome deposits, a distinction is made between hard chrome and bright chrome. Bright chrome is used as a deposit approx. 0.3 μm thick with a nickel or copper/nickel intermediate deposit. In the past, bumpers and molding strips were provided with bright-chrome deposits for decorative reasons. Decorative chrome deposits have become an important element in vehicle design.

Hard-chrome deposits are chrome deposits with a thickness greater than 2 μm . Owing to the high level of hardness of the electrolytically deposited chrome coating, it is ideally suited for use as a wear-protection coating. In the past, hard-chrome deposits were often applied in strong thicknesses and then mechanically reworked. Due to further developments in process engineering, components today are increasingly custom-coated with coating thicknesses ranging from 5 to 10 μm ,

and then used without reworking (e.g. components for fuel injectors).

Tin

Electrolytically deposited tin coatings are mainly used as contact surfaces for plug-in and switch contacts, and as solder contact surfaces. A coating thickness of 2 to 3 μm is ideal for plug-in contacts. Greater coating thicknesses are required for solder applications in order to ensure solderability even after extended storage periods.

Gold

Gold deposits are normally used for contacts subject to stringent requirements. They are characterized by good conductivity, low contact resistance, and good resistance to corrosion and pollutant gases. This ensures contact stability. Hard-gold deposits (with approx. 0.5 % alloying constituents, predominantly cobalt) are harder and more abrasion-resistant than pure-gold deposits and are suitable for contacts subject to mechanical load.

Chemically deposited coatings (no external current)

In contrast to electrodeposits, chemically deposited coatings are characterized by a more uniform coating-thickness distribution on the component, since deposition in a coating bath does not occur under

Table 1: Areas of application for coatings

Coating system	Coating material	Main application
Electrodeposits	Zn, ZnNi, Ni Cr Sn, Ag, Au	Corrosion protection Wear protection Decorative Electrical contacts
Chemical deposits	NiP, NiP dispersions Cu, Sn, Pd, Au	Wear and corrosion protection Electronic applications
Hot-immersion deposits	Zn, Al Sn	Corrosion protection Electrical contacts
Paints (wet and powder-based paints)	Organic polymers Pigments (color particles)	(Decorative) corrosion protection Wear reduction Electrical insulation
PVD and CVD coatings (plasma technology)	TiN, TiCN, TiAlN DLC i-C(WC), a-C:H	Wear protection of tools Reduction of friction and wear of components

the influence of an external electric field. Owing to the slow deposition rate and the costly chemicals in the coating bath, they are more expensive than electroplated liners. The following have become widespread in industry:

- “Chemical nickel” (nickel-phosphorus, NiP) as a corrosion- and wear-protection deposit, used for example in electric fuel pumps or in high-pressure pumps for gasoline direct injection.
- “Chemical copper” and “Chemical tin” in printed-board technology.

Hot-immersion deposits

Hot-immersion coatings are deposited by immersing the substrates in a molten metal bath. Cathodically effective hot-immersion deposits are used to protect low-alloyed steels against corrosion. Zinc, zinc/aluminum alloys and aluminum are used. Coating start material such as sheet and strip is a cheap and widely used practice. However, free punch edges must be accepted here.

Tin and tin-alloy coatings deposited in hot-immersion priming are primarily used as surfaces for plug connectors and as soldering surfaces.

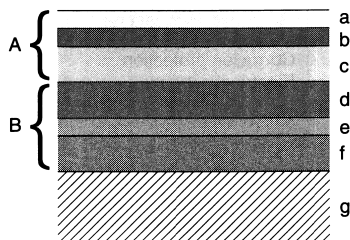
Figure 2: Layered structure of paint coatings on car bodies

- A Visual appearance,
- B Corrosion protection.
- a Top coat, transparent,
- b Top coat, colored,
- c Extender,
- d Cathodic deposition (CD),
- e Phosphate coat,
- f Zinc electrodeposit,
- g Sheet steel.

Paint structure (coats, see Table 2):

Solid: 1, 2, 3a, 3b, 4, 5.

Metallic: 1, 2, 3c, 3d, 4, 5.



SAM0115-1D

Zinc-lamella deposits

Zinc-lamella deposits are coatings based on zinc and aluminum lamellas in an inorganic binder. They are applied by immersion centrifuging or electrostatic spraying and thermal hardening. Zinc-lamella deposits are low-cost corrosion-protection coatings for mass-produced steel parts (e.g. screws/bolts).

Paints

Paints have a wide variation range because of their broad chemical bases and numerous means of application, such as brushing, spraying, and immersing, even with the use of current (CD: cathodic deposition).

Paint coatings can also fulfill a multitude of different functions. In motor vehicles, the primary function fulfilled by paint coatings is corrosion protection accompanied by a decorative effect, but other functions include wear protection, soundproofing or electrical insulation.

The very high demands imposed on corrosion protection and visual appearance over the service life of a vehicle are guaranteed by a complex, elaborate layered structure (Figure 2 and Table 2). The assemblies in the engine compartment, on the other hand, are usually only given one or two coats of paint. Less importance is placed here on decorative properties.

Low-solvent systems, above all water-based paints, are almost always used in automotive engineering. Where possible, powder-based paints and UV-hardening paints which are completely solvent-free are used.

Table 2: Structure of solid-color and metallic coatings

Layer	Layer thickness in μm	Structure	Composition		Solvents	Pigments	Extenders	Additives and SC		Applications
			Binders					Inorganic extenders	Surface-active substances, anti-crater agents, SC 20 %	
1	20 to 25	Cathodic deposition	Epoxy resins Polyurethane		Water, small amounts of water-miscible organic solvents	Inorganic (organic)				Electrophoretic coating
2a	approx. 35 to 50	Extender	Polyester, melamine, urea, and epoxy resins		Aromatic compounds, alcohols	Inorganic and organic	Inorganic solids		E.g. wetting agents, surface-active substances SC 55... > 70 %	PS ESTA-HR
2b	approx. 35 to 50	Water extender	Water-thinnable polyester, polyurethane, and melamine resins		Water, small amounts of water-miscible organic solvents				SC 43 to 50 %	PS ESTA-HR
3a	40 to 50	Solid-color top coat	Alkyd and melamine resins		Esters, aromatic compounds, alcohols		-		E.g. leveling, wetting agents SC 50 to 60 %	PS ESTA-HR
3b	10 to 35 (color-dependent)	Water-thinnable solid-color base coat	Water-thinnable polyester, polyurethane, polyacrylate, melamine resins		Small amounts of water-miscible cosolvents				Wetting agents SC 20 to 40 %	PS ESTA-HR
3c	10 to 15	Metallic, base coat	CAB, polyester, melamine resins		Esters, aromatic compounds	Aluminum and mica particles	-		SC 15 to 30 %	PS ESTA-HR
3d	10 to 15	Water-thinnable metallic base coat	Water-thinnable polyester, polyurethane, and polyacrylate, and melamine resins		Small amounts of water-miscible cosolvents	Aluminum mica particles Organic and inorganic pigments	-		Wetting agents SC 18 to 25 %	PS ESTA-HR
4a	40 to 50	Clear coat conventional	Acrylic and melamine resins		Aromatic compounds, alcohols, esters	-	-		E.g. leveling agents and light stabilizers, SC 45 %	PS ESTA-HR
4b	40 to 50	2C-HS	HS acrylate resin polyisocyanates		Esters, aromatic compounds	-	-		E.g. leveling agents and light stabilizers, SC 58 %	
4c	40 to 50	Powder-slurry clear coat	Urethane-modified epoxy/carboxy system		-	-	-		E.g. light stabilizers, SC 38 %	
4d	40 to 50	Powder clear coat	Acrylate resin		-	-	-		100 %	ESTA PS

Acronyms: ESTA-HR Electrostatic high rotation; SC Solids content; PS Pneumatic spray; 2C-HS 2-components high solid (high levels of non-volatile matter)

PVD and CVD coatings

PVD and CVD coating systems are generally deposited in a vacuum supported by plasma or heat treatment on components and tools (Figure 3). There are two different procedures, depending on whether the coating-forming material is obtained from the solid (PVD, physical vapor deposition) or gaseous (CVD, chemical vapor deposition) phase. Most modern procedures combine the two categories.

Hard-material coatings

Tools are coated with hard material to increase their service life and their performance. A typical example of the tool coatings widely introduced onto the market is the gold-colored titanium nitride (TiN), which is produced, for example, by the cathodic sputtering or arc evaporation of titanium in a reaction with nitrogen. More recent coating systems, such as titanium carbonitride (TiCN) or titanium aluminum nitride (TiAlN), can be used for high-speed cutting or for dry machining.

Superhard materials, such as diamond, are increasingly used to coat carbide tools.

DLC coatings

Protecting components against wear is subject to special conditions. Here, the coating should establish a low coefficient of friction for the components that come into contact with and move relative to each other and minimize the wear for the entire layout. DLC (diamond-like carbon) coatings protect not only the coated component against wear, but also the uncoated opposed body. They have a very low coefficient of friction of 0.1 to 0.2, both in a dry friction pairing against steel and under media. DLC layers are media-resistant and have a corrosion-inhibiting effect. Due to these advantages, DLC coatings are very important in the field of coating components. However, it must be noted that coatings containing hydrogen will degrade in an oxidizing atmosphere at a local temperature of 350 °C and higher. Hard-material coatings, such as TiN, can withstand considerably higher temperatures and are also used in component coating.

Different coating compositions and processes allow materials to be adapted to different wear loads or combinations of abrasive, vibration, sliding, seizure, and adhesive wear. The term DLC therefore describes not a layer but a class of materials within which there exists a multitude of coating variants with all manner of properties and with manufacturer-specific nomenclature.

Plasma-assisted PVD and CVD processes are conducted in such a way that the tempering temperature of the component is not exceeded during coating. This avoids impairing the hardness of the steel base material.

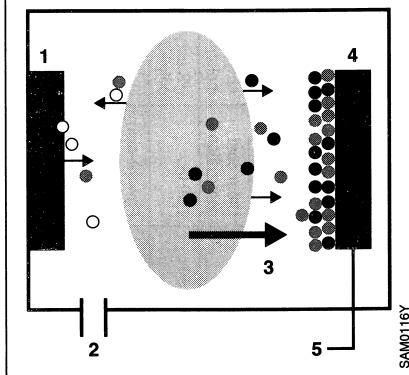
Low-friction carbon coatings containing metal carbide i-C(WC) are electrically conductive and have a microhardness of approx. 1,800 HV at a modulus of elasticity of 150 to 200 GPa.

In comparison, metal-free carbon coatings (a-C:H) offer increased hardness of approx. 3,500 HV and significantly improved wear resistance, but this is ac-

Figure 3: Schematic of plasma-assisted PVD/CVD coating in vacuum chamber

- 1 Cathode with sputter target,
- 2 Gas inlet,
- 3 Intensive ion bombardment,
- 4 Substrate,
- 5 Substrate tension

- Argon,
- Metal/carbon,
- Nitrogen/hydrogen.



accompanied by an increase in brittleness. They are electrically insulating.

The greatest coating hardness and wear resistance for DLC coatings is currently achieved with metal- and hydrogen-free ta-C (tetrahedral amorphous carbon). For all DLC coatings, in view of their high internal compression stress, it is important to ensure a stable connection to the substrate, which is achieved for the most part with a metallic adhesive coating system.

With thicknesses of 2 to 4 μm , diamond-like carbon coatings offer very good wear protection and are especially suitable for precision components subject to high mechanical loads. There is no rework needed after coating.

Thermal spray-coating

In thermal spraying metals or alloys are melted thermally or in a plasma flame and sprayed onto the surface to be coated. There are a multitude of ways in which the procedure is modified in which ceramic surfaces can be created as well as metallic coatings. By choosing the relevant procedure, it is possible to create tightly sealed coatings for example for wear or corrosion protection or porous, permeable coating systems.

Coatings

Diffusion coatings

Surface treatment can be selectively combined with surface hardening by using the diffusion process to thermochemically carburize, carbonitride, or chromate the metal, or treat it with boron or vanadium (see Heat treatment of metallic materials, Thermochemical treatment). The metal can also be oxidized, nitrided, or sulfided without hardening.

Conversion coatings

Conversion coatings are formed not exclusively by the application of a material, but by in part the chemical or electrochemical conversion of the base material.

Browning coatings

Browning coatings consist of thin iron-oxide layers (predominantly Fe_3O_4) which are formed by the oxidation of the iron in steel in an alkaline, aqueous solution containing nitrite at temperatures greater than 100 °C. With subsequent oiling, they offer temporary corrosion protection for low-alloy steels.

Phosphating coatings

Phosphating coatings are formed on steel, galvanized steel, and aluminum in solutions containing phosphoric acid by immersion or spraying. Zinc-phosphate coatings are predominantly used as a wash primer for paint coatings or in combination with anti-corrosion oil. Manganese phosphate serves as a wear-protection coating with anti-seizure properties, and as a wash primer for other coatings to improve component antifriction properties.

Anodized coatings

Anodized coatings are formed by the electrochemical conversion of metal into metal oxide in aqueous electrolytes. Aluminum, magnesium, and titanium can be anodized. Anodized coatings on aluminum materials are widely used for corrosion and wear protection.

Lubricants

Terms and definitions

Lubricants

Lubricants provide mutual insulation for components in a state of relative motion. The lubricant's function is to prevent direct contact between these components, thereby reducing wear and minimizing friction. Lubricants serve as coolants, friction-surface sealants, and corrosion inhibitors, and can also reduce running noise. Lubricants may be solid, consistent, liquid, or gaseous in form. Specific lubricants are selected with reference to design characteristics, materials combinations, the ambient conditions, and the stress factors encountered at the friction surface ([1], [2], [3]).

Additives

Additives are substances mixed into the lubricant in order to improve specific properties. These substances modify either the lubricant's physical characteristics (e.g. viscosity-index improvers, pour-point depressors) or its chemical properties (e.g. oxidation inhibitors, corrosion inhibitors). In addition, the properties of the friction surfaces themselves can be modified by additives which change the friction characteristics (friction modifiers), protect against wear (anti-wear agents), or provide protection against scoring and seizure (extreme-pressure additives). Great care must be exercised in order to ensure that the additives are correctly matched with each other and with the base lubricant.

ATF

ATF (automatic transmission fluids) are special-purpose lubricants specifically formulated to meet stringent requirements for operation in automatic transmissions.

Ash

Ash is the mineral residue which remains after oxide and sulfate incineration (DIN 51575, [4]).

Bleeding

Bleeding is the separation of base oil (lube oil) and the thickener in a lubricating grease (oil separation, DIN 51817 [5]). During this process the lube oil is released to the lubricating surroundings.

Bingham bodies

Bingham bodies are materials whose flow characteristics differ from those of Newtonian fluids (see Rheology, Newtonian fluids).

Fire point and flash point

The lowest temperature (referred to 1,013 hPa) at which a gaseous mineral product initially flashes is known as the flash point. The fire point is the temperature at which it continues to burn for at least 5 s (DIN EN ISO 2592 [6]).

Cloud point

The cloud point is the temperature at which mineral oil becomes opaque due to the formation of paraffin crystals or precipitation of other solids (DIN ISO 23015 [7]).

Detergents

These additives prevent paint- and carbon-like deposits on hot components (e.g. pistons). They serve as cleaning agents which like a tenside contain both hydrophilic and hydrophobic areas.

Dispersants

Dispersants prevent (disperse) sludging and sedimentation, particularly at low temperatures (solid substances are held finely distributed in suspension).

EP lubricants

See High-pressure lubricants (EP, Extreme Pressure).

Flow pressure

The flow pressure is the pressure required to press a consistent lubricant through a standardized test nozzle (Kesternich method, DIN 51805 [8]). The flow pressure is an index of a lubricant's starting flow characteristics, particularly at low temperatures.

Yield point

The yield point is the shear stress at which a substance begins to flow. Above the yield point, the rheological characteristics of plastic substances are the same as those of liquids (DIN 1342-3 [9]).

Friction modifiers

Friction modifiers are polar lubricant additives which reduce friction in the mixed-friction range (see Mixed friction) and increase bearing capacity after adsorption on the surface of the metal. They also inhibit stick-slip behavior.

Gel-type greases

Gel-type greases are lubricants with inorganic gelling agents (e.g. Bentonites, silica gels).

Anti-friction coatings

Anti-friction coatings are solid lubricant combinations which a binding agent holds in place on the friction surfaces (AFC, anti-friction coating).

Graphite

Graphite is a solid lubricant with layer-lattice structure. Graphite provides excellent lubrication when combined with water (e.g. high atmospheric humidity) and in carbon-dioxide atmospheres or when combined with oils. It does not inhibit friction in a vacuum.

Borderline pumping temperature

The borderline pumping temperature is the application limit for oil throughput through engines after cold starting. Below the borderline pumping temperature oil can no longer flow of its own accord to the oil pump, and thus a sufficient supply of oil can no longer be guaranteed.

High-pressure lubricants

High-pressure lubricants (EP, Extreme Pressure) contain additives to enhance load-bearing capacity, to reduce wear, and to reduce scoring (generally provide good performance in steel-to-steel and steel-to-ceramic applications).

Hydro-crack oils

Hydro-crack oils are refined mineral oils with increased viscosity index (VI 130 to 140). They are produced by hydrogenating mineral oils and are thus more thermally stable and have improved viscosity-temperature characteristics.

Induction period

The induction period is the time which elapses before substantial changes occur in a lubricant (e.g. aging of an oil containing an oxidation inhibitor).

Inhibitors

Inhibitors are active system-protecting ingredients which slow down the chemical decomposition of lubricants (e.g. due to oxidation) or protect against corrosion (corrosion inhibitors).

Low-temperature sludge

Low-temperature sludges are the products of oil degradation which form in the engine crankcase due to incomplete combustion and condensate at low engine load. Low-temperature sludge increases wear and can cause engine damage. Modern high-quality engine oils inhibit its formation.

Cold-starting reliability

Cold-starting reliability is a temperature value and based on experience is approx. 5 °C above the borderline pumping temperature.

Consistency

Consistency is a measure of the ease with which lubricating greases and pastes can be deformed (DIN ISO 2137 [10]). A standardized cone is dropped into a flattened grease sample and then the penetration depth is measured in 1/10 mm (Figure 2, see also Penetration and Table 5).

Doped lubricants

Doped lubricants contain additives for improving specific properties (e.g. aging stability, wear protection, corrosion protection, viscosity-temperature characteristics).

High-lubricity oils

High-lubricity oils are lubricating oils with multigrade properties, low cold viscosity, and special anti-friction additives. Extremely low engine friction under all operating conditions reduces fuel consumption.

Longlife engine oils

Longlife engine oils are oils for significantly extended oil-change intervals.

Multigrade oils

Multigrade oils are engine and transmission oils with good resistance to viscosity-temperature change (high viscosity index VI, Figure 1, see Viscosity). They often contain polymers which change their solid structure as a function of temperature. When the oil is cold, the molecules in the oil are "bunched together". As the temperature increases, the molecules stretch and thereby increase the friction between the particles. The viscosity of the oil increases, thereby guaranteeing a stable lubricating film even at high temperatures. Multigrade oils reduce friction and wear, and during cold starting ensure that all the engine components are quickly lubricated. These oils are formulated for year-round use in motor vehicles; their viscosity ratings extend through several SAE grades.

Metal soaps

Metal soaps are reaction products from metals or from their compounds with fatty acids. They are used as thickeners for grease and as friction modifiers.

Mineral oils

Mineral oils are distillates or raffinates produced from petroleum or coal. They consist of numerous hydrocarbons in various chemical compositions. Classification is according to the predominant component: paraffin-based oils (chain-shaped saturated hydrocarbons), naphthene-based oils (closed-chain saturated hydrocarbons, generally with five or six carbon atoms per ring) or aromatic oils (e.g. alkylbenzene). These substances are distinguished by major variations in their respective chemical and physical properties.

Molybdenum disulfide

Molybdenum disulfide (MoS_2) is a solid lubricant with layer-lattice structure. Only low cohesive forces are present between the individual layers, so their mutual displacement is characterized by relatively low shear forces. A reduction in friction is only obtained when MoS_2 is applied in suitable form to the surface of the metal (e.g. in combination with a binder such as MoS_2 anti-friction coating).

Figure 1: Viscosity/temperature curves for single- and multigrade engine oils

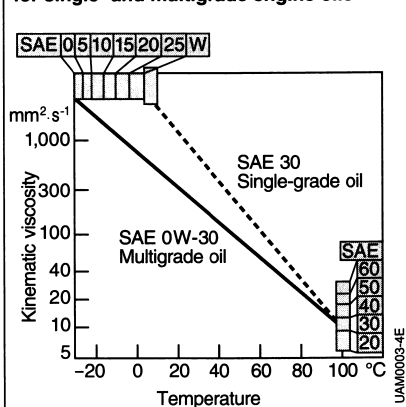
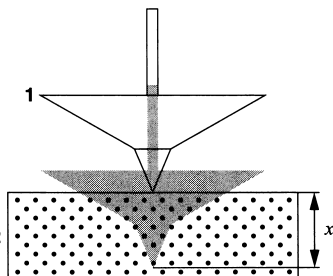


Figure 2: Determining cone penetration

- 1 Standardized cone,
 - 2 Consistent lubricant at defined temperature.
- x Penetration depth.



Penetration

Penetration denotes the depth (in 10^{-1} mm) to which a standardized cone penetrates into a consistent lubricant within a defined period and at a specified temperature (Figure 2). The larger the number, the softer the lubricant (DIN ISO 2137 [10]).

Polar substances

Dipolar molecules are easily adsorbed onto metal surfaces. They enhance adhesion and bearing capacity, thus reducing friction and wear. This category includes, for example, diester oils, ethers, polyglycols, and fatty acids.

Pour point

Pour point is the lowest temperature at which an oil continues to flow when cooled under defined conditions (DIN ISO 3016 [11]).

PTFE

PTFE (polytetrafluoroethylene, Teflon) is a thermoplastic with outstanding properties as a solid lubricant, particularly at very low sliding velocities (< 0.1 m/s). PTFE only becomes brittle below approx. -270 °C. The upper service temperature for use is approx. 260 °C. Above this level, it decomposes with toxic cleavage products.

Rheology

Rheology is the science dealing with the flow characteristics of materials. These are generally represented in the shape of flow curves (Figure 3). In the diagram

shear stress τ is plotted against shear rate $\dot{\gamma}$.

Shear stress and shear rate

Shear stress is defined as the shear force per shear area, shear rate as the quotient of velocity v and gap height h (Figure 4).

Shear stress $\tau = F/A$ (in $\text{N/m}^2 = \text{Pa}$),
 F shear force, A area vs. shear rate.

Shear rate $\dot{\gamma} = v/y$ (in s^{-1}),
 v velocity, y thickness of lubricating film.

The flow curves are usually measured in a plate-cone measuring system. The sample to be measured is located between the plate and the cone; the torque is proportional to the shear stress and the rotational speed proportional to the shear rate.

Shear stress τ can be illustrated using the two-plate model (Figure 4). The fluid is located between the two plates. The bottom plate is fixed, the top plate moves with shear area A by means of shear force F .

Shear viscosity

Shear viscosity η is in liquids of ideal viscosity and constant temperature the ratio of shear stress τ to shear rate $\dot{\gamma}$:

$$\eta = \frac{\tau}{\dot{\gamma}}$$

The term "dynamic viscosity" is also used for η . The former unit centiPoise (cP) is equal to the unit $\text{mPa} \cdot \text{s}$.

Figure 3: Flow curves

- 1 Rheopexic, 2 Thixotropic, 3 Newtonian, 4 Plastic, 5 Dilatant, 6 Intrinsically viscous, 7 Yield point.

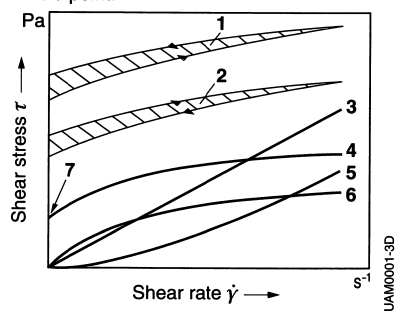
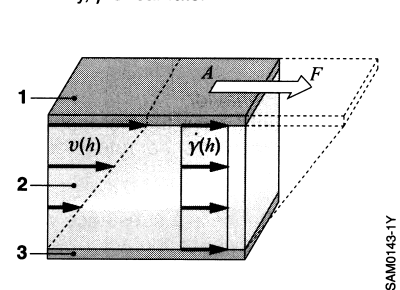


Figure 4: Velocity distribution and shear rate in the lubricating gap

- 1 Moving plate, 2 Fluid, 3 Fixed plate.
 A Shear surface, F Shear force, h Height,
 v Velocity, $\dot{\gamma}$ Shear rate.



Midpoint viscosity

Each viscosity grade in accordance with DIN ISO 3448 [12] is defined by a viscosity range at a temperature of 40 °C (see Table 1). The permissible limit is $\pm 10\%$. Thus, for example, viscosity grade ISO VG 10 has the limits 9 and 11 mm²/s at a midpoint viscosity of 10 mm²/s.

HTHS viscosity

The HTHS viscosity (High Temperature High Shear) specifies the shear viscosity at a temperature of 150 °C and a shear rate of 10^6 s^{-1} .

Kinematic viscosity

The kinematic viscosity ν is usually determined in vertical, narrow glass tubes. The time the fluid needs at a defined temperature to flow through a specific section of the tube is measured. If density ρ is known, it is also possible to calculate the shear viscosity η from it.

$$\nu = \frac{\eta}{\rho}, \quad (\nu \text{ in mm}^2/\text{s}; \rho \text{ in kg/m}^3).$$

The former unit “centiStokes” (cSt) is equal to the unit mm²/s.

Newtonian fluids

Newtonian fluids display a linear relationship between shear stress τ and shear rate $\dot{\gamma}$ in the shape of a straight line through zero, with the slope increasing as a function of viscosity (Figure 3).

All materials not characterized by this kind of flow behavior are classified as non-Newtonian fluids.

Intrinsically viscous flow behavior

Intrinsically viscous flow behavior is the capacity of a fluid to show decreasing viscosities with increasing shear rate (Figure 3, e.g. liquid greases, multigrade oils with viscosity-index improvers).

Dilatant flow behavior

Dilatant flow behavior is the increase in viscosity with increasing shear rate (Figure 3).

Plastic flow behavior

A plastic substance is a substance whose rheological behavior is characterized by a yield point (according to DIN 1342-1 [13]). Plastic flow behavior is the formability

of an intrinsically viscous fluid supplemented by yield point (Figure 3, e.g. lubricating greases).

Thixotropy

Thixotropy is a characteristic of those non-Newtonian fluids that display an decrease in viscosity proportional to shear time, and only gradually recover their original viscosity once shearing ceases (Figure 3).

Rheopexy

Rheopexy is a characteristic of those non-Newtonian fluids that display an increase in viscosity proportional to shear time, and only gradually recover their original viscosity once shearing ceases (Figure 3).

Table 1: Viscosity grades for industrial lubricating oils acc. to ISO 3448 [12]

ISO viscosity grade	Medium viscosity at 40 °C in mm ² /s	Limits of kinematic viscosity at 40 °C in mm ² /s	
		min.	max.
ISO VG 2	2.2	1.98	2.42
ISO VG 3	3.2	2.88	3.52
ISO VG 5	4.6	4.14	5.06
ISO VG 7	6.8	6.12	7.48
ISO VG 10	10	9.00	11.0
ISO VG 15	15	13.5	16.5
ISO VG 22	22	19.8	24.2
ISO VG 32	32	28.8	35.2
ISO VG 46	46	41.4	50.6
ISO VG 68	68	61.2	74.8
ISO VG 100	100	90.0	110
ISO VG 150	150	135	165
ISO VG 220	220	198	242
ISO VG 320	320	288	352
ISO VG 460	460	414	506
ISO VG 680	680	612	748
ISO VG 1000	1,000	900	1,100
ISO VG 1500	1,500	1,350	1,650
ISO VG 2200	2,200	1,980	2,420
ISO VG 3200	3,200	2,880	3,520

Dropping point

The dropping point is the temperature at which a lubricating grease attains a specified viscosity under specified test conditions (DIN ISO 2176 [14]).

Stribeck curve

The Stribeck curve portrays friction levels between two liquid- or grease-lubricated tribological systems separated by a narrowing gap (e.g. lubricated plain or ball bearings) as a function of sliding velocity (Figure 5). As the relative speed increases, so too does the hydrodynamic pressure in friction contact.

Solid-body friction

In the case of solid-body friction, the height of the lubricating film is lower than that of the roughness protrusions in the material's surface. This causes wear.

Mixed friction

In the case of mixed friction, the height of the lubricating film is approximately equal to that of the roughness protrusions. This still means increased wear because the roughness protrusions are in contact with each other.

Hydrodynamics

In the case of hydrodynamics, there is complete separation between the primary body and the opposed body, e.g. when aquaplaning (virtually wear-free condition).

Synthetic oil

Synthetic oil is manufactured in a process of chemical synthesis from smaller molecules. For example, synthetic hydrocarbons in the form of poly- α -olefins, which are manufactured by the polymerization and subsequent hydrogenation of ethylene. Further synthetic oils are polyglycols, diester oils, silicone oils, and perfluoropolyether oils.

Viscosity

Viscosity is a measure of the internal friction of substances (DIN 1342 [13], DIN EN ISO 3104 [15]). It is caused by the resistance (internal friction) with which the substance's molecules oppose displacement forces (see Rheology). As

the temperature decreases, the viscosity increases and the oil becomes more viscous. At low temperatures the viscosity must not be too high so that the oil in the bearings does not exert too excessive a resistance on the rotational motion of the engine or transmission. At high temperatures the oil must still be sufficiently viscous to maintain the lubricating film.

Viscosity index

The viscosity index (VI) is a mathematically derived number expressing the change in a mineral-oil product's viscosity relative to its temperature. The greater the viscosity index, the lower the effect of temperature on the viscosity (DIN ISO 2909 [16]).

Viscosity grades

Oils are classified within specific viscosity ranges in viscosity grades:

- ISO viscosity grades (DIN ISO 3448 [12]): see Table 1.
- SAE viscosity grades (SAE J300 [17], SAE J306 [18]): see Table 2 and Table 3.

Worked penetration

Worked penetration is the penetration of a grease sample after it has been processed in a grease kneader (DIN ISO 2137 [19]).

Figure 5: Stribeck curve

R Surface roughness, F_N Normal force, d Distance between primary and opposed body.
Range a: Solid-body friction, high wear.
Range b: Mixed friction, moderate wear.
Range c: Hydrodynamics, no abrasive wear.

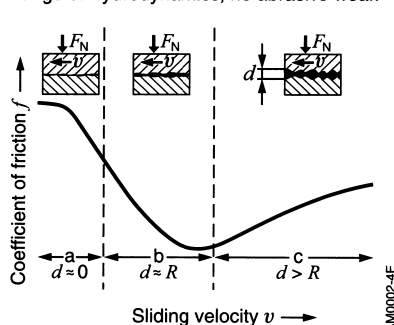


Table 2: Extract from SAE viscosity grades for transmission oils (SAE-J306 revision June 2005)

SAE vis- cosity grade	Maximum temperature [°C] for dynamic viscosity at 150,000 mPa·s (ASTM D 2983) [20])	Kinematic vis- cosity [mm ² /s] at 100 °C (ASTM D 445 [21])	
		min.	max.
70 W	−55	4.1	–
75 W	−40	4.1	–
80 W	−26	7.0	–
85 W	−12	11.0	–
80	–	7.0	<11.0
85	–	11.0	<13.5
90	–	13.5	<18.5
110	–	18.5	<24.0
140	–	24.0	<32.5
190	–	32.5	<41.0
250	–	41.0	–

Engine oils

Engine oils are employed primarily to lubricate contiguous components in relative motion within the internal-combustion engine. The oil also removes heat generated by friction, carries abraded particles away from the friction surface, washes out contaminants, holds them in suspension, and protects metals against corrosion. The most common engine oils are mineral oils treated with additives (HD oils: Heavy Duty for high load, extreme operating conditions). Higher stress-resistance requirements combined with extended oil-change intervals have led to widespread application of fully and semi-synthetic oils (mixture of synthetic oil and mineral oil), e.g. hydro-crack oils. The quality of an engine oil is determined by its origin, the refining processes used on the mineral oil (except in the case of synthetic oils) and the additive composition.

Table 3: SAE viscosity grades for engine and transmission oils (SAE J300, January 2015)

SAE vis- cosity grade	Viscosity (ASTM D 5293 [22])	Limit pumping viscosity (ASTM D 4684 [23]) with no yield point	Kinematic viscosity (ASTM D 445 [21])		Viscosity under high shear ¹ (ASTM D 4683 [24], CEC L-36-A-90 [25], ASTM D 4741 [26] or ASTM D 5481 [27]) mPa·s at 150 °C and $\dot{\gamma} = 10^6 \text{ s}^{-1}$ min.
	mPa·s max.	mPa·s max.	mm ² /s at 100 °C min.	max.	
0 W	6,200 at −35 °C	60,000 at −40 °C	3.8	–	–
5 W	6,600 at −30 °C	60,000 at −35 °C	3.8	–	–
10 W	7,000 at −25 °C	60,000 at −30 °C	4.1	–	–
15 W	7,000 at −20 °C	60,000 at −25 °C	5.6	–	–
20 W	9,500 at −15 °C	60,000 at −20 °C	5.6	–	–
25 W	13,000 at −10 °C	60,000 at −15 °C	9.3	–	–
8	–	–	4	<6.1	1.7
12	–	–	5	<7.1	2.0
16	–	–	6.1	<8.2	2.3
20	–	–	6.9	<9.3	2.6
30	–	–	9.3	<12.5	2.9
40	–	–	12.5	<16.3	3.5 (0W–40, 5W–40, 10W–40)
40	–	–	12.5	<16.3	3.7 (15W–40, 20W–40, 25W–40, 40 monograde)
50	–	–	16.3	<21.9	3.7
60	–	–	21.9	<26.1	3.7

¹ Also called HTHS (High Temperature High Shear) viscosity.

Additives are classified according to their respective functions:

- Viscosity-index improvers
- Pour-point improvers
- Oxidation and corrosion inhibitors
- Detergent and dispersant additives
- Extreme-pressure (EP) additives
- Friction modifiers
- Anti-foaming agents

Oil is subjected to considerable thermal and mechanical stresses in the IC engine. The data on the physical properties of engine oils provide information on their application limits (SAE viscosity grades), but are not indicative of their other performance characteristics. Therefore, there are several different test procedures for evaluating engine oils:

- ACEA (Association des Constructeurs Européens de l'Automobile) standards, replaced the CCMC standards (Comité des Constructeurs d'Automobiles du Marché Commun) at the beginning of 1996
- API classification (American Petroleum Institute)
- MIL specifications (Military)
- Company specifications (e.g. ILSAC, International Lubricants Standardization and Approval Committee).

The approval criteria include the following:

- Sulfate-ash content
- Zinc content
- Engine type (diesel or spark-ignition engines, naturally aspirated or forced-induction engines)
- Load on power-transmission components and bearings
- Wear-protection properties
- Oil operating temperature (in oil pan)
- Combustion residue and chemical stress exerted on the oil by acidic combustion products
- The oil's detergent and residue-scavenging properties
- Its suitability for use with gasket and sealing materials.

ACEA (CCMC) Specifications

Engine oils for gasoline engines

- A1: Special high-lubricity oils with reduced viscosity at high temperatures and high shear for reducing viscous friction.

- A2: Conventional and high-lubricity engine oils without any restriction on viscosity grades. Higher requirements than CCMC G4 and API SH, withdrawn.
- A3: Oils of this category meet higher requirements than A2 and CCMC G4 and G5.
- A5: Improved "fuel economy" properties compared with A3 occasioned by lower viscosity including improved additive composition. Only for use in engines which have been specifically designed for this purpose.

Engine oils for passenger-car diesel engines

- B1: Corresponding to A1 for low friction losses and, consequently, reduced fuel consumption.
- B2: Conventional and high-lubricity engine oils compliant with the current minimum requirements (higher than those of CCMC PD2), withdrawn.
- B3: Exceeds B2.
- B4: Corresponds to B2, particularly suitable for VW TDI engines.
- B5: Oils surpass B3 and B4, improved "fuel economy" properties, also satisfies VW 50600 and 50601. Only for use in engines which have been specifically designed for this purpose.

Engine oils for passenger-car diesel engines with particulate filters

- C1: Since 2004, sulfate-ash content max. 0.5 %. Lowered HTHS (Ford).
- C2: Since 2004, sulfate-ash content max. 0.8 %. With HTHS ≥ 2.9 mPa·s (Peugeot).
- C3: Since 2004, sulfate-ash content max. 0.8 %. With HTHS ≥ 3.5 mPa·s (MB and BMW).
- C4: Since 2007, sulfate-ash content max. 0.5 %. With HTMS ≥ 3.5 mPa·s.

Engine oils for commercial-vehicle diesel engines

- E1: Oils for naturally aspirated and turbocharged engines with normal intervals between oil changes; withdrawn.
- E2: Derived from MB Specification Sheet 228.1. Primarily for engine designs predating the Euro II standard.
- E3: For Euro II engines, derived from MB Specification Sheet 228.3. In com-

- parison with the predecessor category CCMC D5, these oils evince a significant improvement in soot dispersion capability and a much reduced tendency to thicken; withdrawn.
- E4: Diesel engines with Euro I to Euro III standards and high requirements, particularly for extended intervals between oil changes (according to manufacturer's specification). Based to a large extent on MB Specification Sheet 228.5.
 - E5: For Euro III engines, reduced ash content (withdrawn).
 - E6: For EGR engines (exhaust-gas recirculation) with/without diesel particulate filters and SCR-NO_x engines (Selective Catalytic Reduction, i.e. reduction of NO_x to N₂). Recommendation for engines with diesel particulate filters and operation with sulfur-free fuel, sulfate ash less than 1 % by mass.
 - E7: For engines without diesel particulate filters of most EGR engines and of most SCR-NO_x engines, sulfate ash max. 2 % by mass.
 - E9: For engines with/without particulate filters of most EGR and SCR-NO_x engines, sulfate ash max. 1 % by mass.

Example of a designation

According to the grade specification, a numerical code is usually supplemented.

Example: An A3/B3-04 is an engine oil for spark-ignition engines (Grade A) and diesel engines (Grade B) of quality category 3, tested accord to the ACEA classification grade issued in 2004.

API classification grades

- S Grades (Service) for gasoline engines.
- C Grades (Commercial) for diesel engines.
- SF: For engines produced in the 1980s; withdrawn.
- SG: Applicable since 1988, with more stringent sludge test, improved oxidation stability and wear protection.
- SH: Since mid-1993, corresponds to API SG quality level, but with more stringent process requirements for oil grade testing.
- SJ: Since October 1996, more tests than API SH.

- SL: Since 2001, lower oil consumption compared with SJ, lower volatility, improved engine cleanness, greater resistance to aging.
- SM: Since 2004, for gasoline and light diesel engines with improved wear protection, higher aging stability, improved pumpability, and also used oil.
- SN: Since October 2010, compared with SM improvements in piston cleanliness, sludging, and exhaust-gas treatment capability. Elastomer compatibility is also defined.
- CC: Engine oils for non-turbocharged diesel engines with low stress factors; withdrawn.
- CD: Engine oils for non-turbocharged and turbocharged diesel engines, replaced by API CF in 1994.
- CD-2: API CD requirements, plus additional requirements relevant to two-stroke diesel engines.
- CF-2: Oils with special two-stroke properties (since 1994).
- CE: Oils with CD performance characteristics, with supplementary test operation in US Mack and Cummins engines.
- CF: Replaced API CD in 1994. Specially for indirect injection, even if the sulfur content of the fuel is greater than 0.5 %.
- CF-4: As API CE, but with more stringent test procedure in single-cylinder Caterpillar turbo-diesel engines.
- CG-4: For diesel engines operating under very stringent conditions. Exceeds API CD and CE. Fuel sulfur content less than 0.5 %. Required for engines compliant with post-1994 emission-control legislation.
- CH4: Modern commercial-vehicle engine oil since 1998. Surpasses CG-4 in standards of wear, soot and viscosity. Longer oil-change intervals.
- CI-4: Since 2002 for high-speed four-stroke engines which can only still meet future exhaust-emission legislation with EGR. Suitable for sulfur content greater than 0.5 %.
- CI-4 plus: As CI-4, but with improved soot transport and higher requirements with regard to viscosity increase.
- CJ-4: Valid since 2006. For highway vehicles which must satisfy the USA

2007 emissions standards with diesel fuel (below 500 ppm sulfur content). Suitable for particulate-filter and NO_x reduction catalytic converters.

- CK-4: Improved protection for new diesel engines with among others increased aging stability, shear stability, piston cleanliness. Start 12/2016
- FA-4: For new diesel engines of the 2016/17 generation with low HTHS (2.9 to 3.2 mPa · s at 150 °C) among others to meet the required emission stipulations in the USA.

ILSAC

ILSAC (International Lubricants Standardization and Approval Committee) is a joint standard of General Motors, Daimler, the Japanese Automobile Manufacturers Association, and the American Engine Manufacturers Association.

ILSAC GF-3

In addition to API-SL, the standard requires a fuel-economy test.

ILSAC GF-4

Allocation to API-SM.

ILSAC GF-5

This standard has been in force since October 2010 with the aim of increased fuel reduction, elastomer compatibility, and improved protection of the exhaust-gas treatment systems (comparable with API-SN).

ILSAC GF-6

New specification with the aim of fuel economy for gasoline direct injection and turbocharged engines, expected around 2018. GF-6 A with conventional viscosities, GF-6B with viscosities ≤ SAE 0W16.

SAE viscosity grades

(SAE J300 [17], SAE J306 [18])

The SAE (Society of Automotive Engineers) grades are the internationally accepted standard for defining viscosity. It provides information on the temperature range in which the oils are used. The standard provides no information on the quality of the oil.

A distinction is drawn between single-grade and multigrade oils. Sin-

gle-grade oils are designated for example by SAE 30. Low code numbers denote thin-bodies oils, higher code numbers denote more viscous oils.

Multigrade oils are the type in widespread use today. Two series are employed for the designation (see Table 2 and Table 3), where the letter "W" (Winter) is used to define specific cold-flow properties. These oils are also suitable for low temperatures. The viscosity grades including the letter "W" are rated according to maximum cold viscosity, maximum viscosity pumping temperature and the minimum viscosity at 100 °C.

Viscosity grades without the "W" are rated only according to viscosity at 100 °C.

Viscosity specifications at higher temperature (150 °C) and high shear stress (HTHS viscosity, see Table 3) are more practical. This value acquired particular significance with the introduction of oils with reduced high-temperature viscosity (below 3.5 mPa · s). These oils may only be used if manufacturer approvals exist.

A multigrade oil is designated for example by SAE 0W-30. This means that it has at -35 °C a viscosity of max. 6,200 mPa · s and satisfies the specification of an SAE 30 oil with a kinematic viscosity in the range of 9.3 to 12.5 mm²/s at 100 °C and an HTHS viscosity of min. 2.9 mPa · s at 150 °C (Table 3, see also Figure 1).

Motorcycle oil

This essentially is no different from oil for automobiles. However, the stress for example with regard to shear load and specific surface pressure is if necessary higher and only the oils recommended by the manufacturer may be used.

Oil dilution

When vehicles are operated in short-distance service, unburned fuel can mix with engine oil (oil dilution, oil proliferation). This reduces the operating viscosity and increased wear can occur under mixed-friction conditions. In winter the presence of diesel fuel in engine oil can also cause paraffin-crystal precipitation, which compromises delivery (e.g. filter blockage) and can even cause deficient-lubrication conditions.

Transmission oils

Specifications for transmission oils are defined by the type of transmission and the stresses to which it is subjected throughout the entire range of operating conditions. The requirements (high pressure resistance, high viscosity stability relative to temperature, high resistance to aging, good anti-foaming properties, compatibility with gaskets and seals) can only be satisfied by lubricants treated with special additives. In contrast to engine oils, transmission oils contain no, or only minimal, detergent additives, considerably fewer basic constituents, and mostly no viscosity-index improvers. Most of them would be sheared and thus become inactive. The use of unsuitable and qualitatively inferior oils results typically in damage to bearings and gear-tooth flanks.

The viscosity must also suit the specific application. Viscosity grades for vehicular transmissions are defined in SAE J 306 [18] and SAE J300 [17] (Table 2 and Table 3). Transmission oils with roughly the same viscosity compared with engine oils are identified by higher code numbers so that they can be clearly distinguished from engine oils.

Synthesized oils are being increasingly used to meet special requirements (e.g. poly- α -olefins). Advantages over standard mineral oils include superior temperature-viscosity properties and increased aging resistance.

API classifications of transmission oils

- GL-1...GL-3: Outdated, no longer of practical significance.
- GL-4: Transmission lubricants for moderate-stressed hypoid-gear transmissions, and for transmissions which operate at extreme speeds and impact loads, high rotational speeds and low torques, or low rotational speeds and high torques.
- GL-5: Transmission lubricants for high-stressed hypoid-gear transmissions in passenger cars and in other vehicles where they are exposed to impact loads at high rotational speeds, and at high rotational speeds and low torques, or low rotational speeds and high torques.
- MT-1: For non-synchromesh manual transmissions in US trucks.

Many truck and component manufacturers have drawn up their own specifications and no longer rely on API classification grades.

Lubricants for automatic transmissions

Automatic transmissions differ from their manual-shifted counterparts in the way they transfer torque; non-positive mechanical and hydrodynamic force transfer is supplemented by friction coupling arrangements. Thus, the friction response of automatic transmission fluids (ATF) is extremely important. Applications are basically classified according to the friction characteristics.

General Motors

- Superseded grades: Type A, Suffix A, DEXRON, DEXRON B, DEXRON II C, DEXRON II D.
- DEXRON II E: Valid through 1994.
- DEXRON III F/G: Valid since 1994 and 1997. Features more stringent requirements for oxidation stability and consistency in frictional coefficient.
- DEXRON III H: Valid since 2005, but outdated since 2007.
- DEXRON VI: Valid since 2006.

Other manufacturers

(among others. Ford, MB, MAN, Mack, Scania, ZF): In accordance with fluids and lubricants specifications.

Lubricating oils

Lubricating oils comprise the components of basic oil and additive. The additives improve the properties of the basic oils, e.g. with regard to oxidation stability, corrosion protection, protection against scoring and seizure, or viscosity-temperature characteristics. In addition, system properties such as friction and wear are optimized in the desired direction.

There are a wide range of designations in the form of letters and numbers (see among others DIN 51502 [28]) for the most varied applications. For example, hydraulic fluids:

- HL: Mineral-oil-based hydraulic fluid with additives for improving corrosion protection and resistance to aging.
- HLP: As HL, but with extra additives against scoring and seizure.
- HVLP: As HLP, with extra viscosity-index improver.

Lubricating greases

Lubricating greases are thickened lubricating oils. A great advantage that greases enjoy over oil is that they do not drain from the friction surfaces. Complicated measures to seal them in place are therefore unnecessary (e.g. application in wheel bearings pumps, alternators, wind-shield-wiper motors, servo motors).

Composition

Table 4 provides a general overview of the components in a consistent lubricating grease as blended from three basic components – base oil, thickener, and additive.

Mineral oils are usually employed as the basic oil component, although full-synthetic oils have recently become more common as a replacement (e.g. due to more stringent requirements for aging stability, cold-flow properties, viscosity-temperature characteristics).

Thickeners are used as a binder for the base oil; metal soaps being generally employed (Figure 6). They bind the oil in a sponge-like soap structure (micelle) by means of inclusions and physical Van der Waals forces. The higher the proportion of thickener (depends on the type of thickener) in the grease, the less the penetration and the higher the NLGI grade

Table 4: Composition of lubricating greases

Whatever the friction pairing, the large number of lubricant components can be used to develop a high-performance lubricant.

Base oils	Thickeners	Additives
Mineral oils	Metal soaps of metals	Oxidation inhibitors
– paraffinic	Li, Na, Ca, Ba, Al	Fe, Cu ions, complexing agents
– naphthenic	Normal soaps	Corrosion inhibitors
– aromatic	(soaps with a carbonic acid, e.g. stearic acid)	High-pressure additives (EP additives)
Poly- α -olefins	Hydroxy soaps	Wear-protection additives (anti-wear additives)
Alkyl aromatics	(soaps with an additional hydroxide group, e.g. 12-hydroxystearic acid)	Friction reducers (friction modifiers)
Diester oils	Complex soaps	Adhesion improvers
Polyhydric alcohol	(e.g. Ca soaps with a short- and a long-chain carbonic acid)	Detergents, dispersants
Silicones	Polyureas	VI improvers
Phenyl ether oils	PTFE	Solid lubricants
Perfluoropolyethers	PE	
	Bentonites	
	Silica gels	

(National Lubricating Grease Institute, US standard, see Table 5).

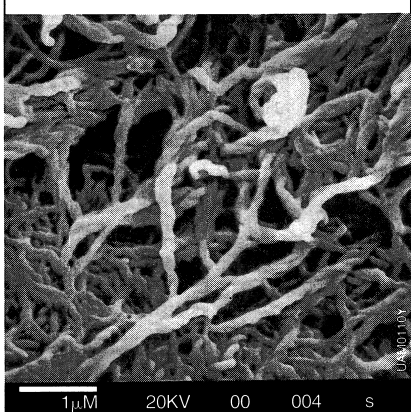
The additives serve to modify physical and chemical properties of the lubricating grease to achieve specific objectives (such as improvement of anti-oxidation properties, increased protection against scoring and seizure, and reduction of friction and wear).

Solid lubricants (e.g. MoS₂) are also added to lubricating greases (for instance for lubricating constant-velocity joints in motor vehicles).

Table 5. Consistency grades for lubricating greases (DIN 51818 [29])

NLGI grade	Worked penetration acc. to DIN ISO 2137 [19] in units of 0.1 mm
000	445...475
00	400...430
0	355...385
1	310...340
2	265...295
3	220...250
4	175...205
5	130...160
6	85 to 150 (unworked penetration)

Figure 6: Photo of lithium soap taken by a scanning electron microscope
The oil is retained between the twisted soap fibrils.



Selection

Specific lubricating greases are selected with reference to their physical characteristics and their effects on the friction surface, and to minimize interaction between the grease and the contact materials. Mutually antagonistic effects with polymers are, for example: formation of stress cracks (Figure 7), consistency changes, polymer degradation, swelling, shrinkage, and brittleness.

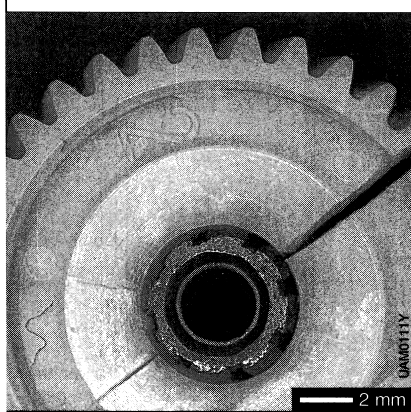
Thus, for example, mineral-oil greases and greases based on synthetic hydrocarbons should not come in contact with elastomers used together with brake fluid (polyglycol base) (e.g. substantial swelling of EPDM elastomers).

In addition, lubricating greases with varying compositions should not be mixed (changes in physical properties, grease liquefaction due to drop-point reduction).

Stress

Thermal and mechanical stresses result in chemical and/or physical changes which may have a detrimental effect on the function of the entire tribological system (Figure 8). Oxidation, for example, results in acidification, which can trigger corrosion on metal surfaces or stress cracking on some plastics. In the event of excessive thermal load, polymerization can cause the lubricant to solidify.

Figure 7: Stress cracks on a gearwheel made of polyoximethylene (POM), caused by poly- α -olefin (PAO)



Every chemical change automatically causes a change in physical properties. These include the rheological properties as well as changes in the viscosity-temperature characteristics or the dropping point. A marked lowering of the dropping point would result in the lubricant flowing away from the friction surface, even at moderate heat.

It is particularly important to remember that metals such as iron or copper (or metals containing copper such as bronze or brass) catalyze the oxidation of a lubricant, i.e. oxidation occurs much more quickly than without catalyst contact. Oxidation quickly renders the lubricity of grease insufficient. Often the soap structure decomposes, the grease then becomes oily, flows away from the friction surface, or hardens due to polymerization.

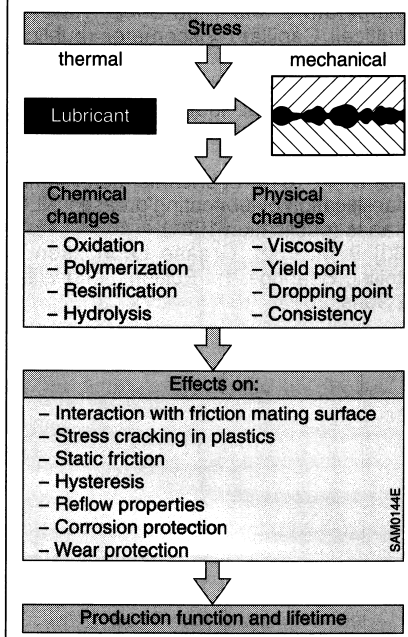
By correctly matching the lubricating grease and tribological system while taking into account the load and interaction, it is possible to enhance the performance potential of products substantially with

sliding-contact opposed parts (e.g. transmission, friction or roller bearings, actuator and control systems) [30].

Lubricants for high voltage requirements

Stronger electric alternating fields than previously occur with increasing operating voltage in alternators and electric motors. This can give rise for example in the ball bearings of motors or alternators to electric discharges between the rolling elements and the raceway. Sparks flash over and can cause tiny areas of the metal to melt. This is called electrical pitting and can cause premature bearing failure. The energy of these discharges will in future be potentially greater if power density and vehicle system voltage increase, e.g. 400 to 1000 V in electric and hybrid vehicles. Without suitable countermeasures there is a high risk of premature bearing damage, initially discernible from increased bearing noise. By reducing the specific resistance of the roller-bearing greases by magnitudes of around 5, for example with ionic liquids, it is possible to reduce dramatically the number and the energy of the discharges.

Figure 8: Stress exerted on lubricant and the resulting effects



References

- [1] T. Mang, W. Dresel: Lubricants and Lubrication. 2nd Edition, Wiley-VCH Verlag GmbH, 2007.
- [2] Thomas Mezger: Das Rheologie-Handbuch. 4th Edition, Vincentz-Network, 2012.
- [3] Wilfried J. Bartz: Schmierfette. 2nd Ed., Springer-Verlag, 2016.
- [4] DIN 51575: Testing of mineral oils – Determination of sulfated ash (publication 2016).
- [5] DIN 51817: Testing of lubricants – Determination of oil separation from greases under static conditions (publication 2014).
- [6] DIN EN ISO 2592: Petroleum products – Determination of flash and fire points – Cleveland open cup method (ISO 2592:2000); German version EN ISO 2592:2002.
- [7] DIN EN 23015: Petroleum products; determination of cloud point (ISO 3015:1992); German version EN 23015:1994.

- [8] DIN 51805: Testing of lubricants; determination of flow pressure of lubricating greases, Kesternich method (publication 1974).
- [9] DIN 1342-3: Viscosity – Part 3: Non-newtonian liquids (publication 2003).
- [10] DIN ISO 2137: Petroleum products – Lubricating grease and petrolatum – Determination of cone penetration (ISO 2137:1985) (publication 1997).
- [11] DIN ISO 3016: Petroleum oils; determination of pour point (publication 1982).
- [12] DIN ISO 3448: Industrial liquid lubricants – ISO viscosity classification (ISO 3448:1992) (publication 2010).
- [13] DIN 1342-1: Viscosity – Part 1: Rheological concepts (publication 2003).
- [14] DIN ISO 2176: Petroleum products – Lubricating grease – Determination of dropping point (ISO 2176:1995) (publication 1997).
- [15] DIN EN ISO 3104: Petroleum products – Transparent and opaque liquids – Determination of kinematic viscosity and calculation of dynamic viscosity (ISO 3104:1994 + Cor. 1:1997); German version EN ISO 3104:1996 + AC:1999.
- [16] DIN ISO 2909: Petroleum products – Calculation of viscosity index from kinematic viscosity (ISO 2909:2002) (publication 2004).
- [17] SAE J300: Engine Oil Viscosity Classification (publication 2015).
- [18] SAE J306: Automotive Gear Lubricant Viscosity Classification (publication 2005).
- [19] DIN ISO 2137: Petroleum products – Lubricating grease and petrolatum – Determination of cone penetration (ISO 2137:1985) (publication 1997).
- [20] ASTM D 2983: Standard Test Method for Low-Temperature Viscosity of Lubricants Measured by Brookfield Viscometer (publication 2015).
- [21] ASTM D 445a: Standard Test Method for Kinematic Viscosity of Transparent and Opaque Liquids (and Calculation of Dynamic Viscosity) (publication 2015).
- [22] ASTM D 5293: Standard Test Method for Apparent Viscosity of Engine Oils and Base Stocks Between -10 and -35 °C Using Cold-Cranking Simulator (publication 2015).
- [23] ASTM D 4684: Standard Test Method for Determination of Yield Stress and Apparent Viscosity of Engine Oils at Low Temperature (publication 2014).
- [24] ASTM D 4683: Standard Test Method for Measuring Viscosity of New and Used Engine Oils at High Shear Rate and High Temperature by Tapered Bearing Simulator Viscometer at 150 °C (publication 2013).
- [25] CEC L-36-A-90: European standard test for HTHS viscosity, technically identical to ASTM 4741.
- [26] ASTM D 4741: Standard Test Method for Measuring Viscosity at High Temperature and High Shear Rate by Tapered-Plug Viscometer (publication 2013).
- [27] ASTM D 5481: Standard Test Method for Measuring Apparent Viscosity at High-Temperature and High-Shear Rate by Multicell Capillary Viscometer (publication 2013).
- [28] DIN 51502: Designation of lubricants and marking of lubricant containers, equipment and lubricating points (publication 1990).
- [29] DIN 51818: Lubricants; consistency classification of lubricating greases; NLGI grades (publication 1981).
- [30] P. M. Lugt: Grease Lubrication in Rolling Bearings. 1st Edition, John Wiley & Sons, 2013.

Coolants for internal-combustion engines

Requirements

When a fuel is combusted in the engine, the supplied energy is converted in the proportions of approx. 1/3 into kinetic energy and approx. 2/3 into heat. Roughly 50 % of this heat escapes via the exhaust system, the remainder is utilized or dissipated in the engine cooling circuit via the coolant. The coolant must for this purpose satisfy a series of important requirements:

- optimal heat-transfer properties,
- high heat capacity,
- low evaporative losses,
- good antifreeze properties,
- corrosion, erosion and cavitation protection of all the cooling-circuit materials,
- compatibility of elastomers, plastics, and coatings,
- avoidance of deposits (“fouling”) which result in blockages,
- low maintenance requirement, long service life, and easy handling,
- environmental compatibility.

Composition of coolants

Customary coolants for engine cooling circuits consist of a mixture of water with a radiator protection agent approved by the automobile and engine manufacturers. The volume percentage of the radiator protection agent is for the most part between 40 % and 50 %.

Water

Owing to its high specific heat capacity and its correspondingly substantial thermal-absorption capacity, water is a very good coolant. Water from the main is usually used. Since the quality of such water can vary dramatically, the requirements with regard to the water used are specified by most automobile and engine manufacturers, and by radiator manufacturers (Table 1). A higher water ion content results in faster inhibitor consumption and possible after-effects such as, for example, flocculation (coagulation) of low-solubility reaction products, corrosion of circuit components, through to blockage of cooling ducts.

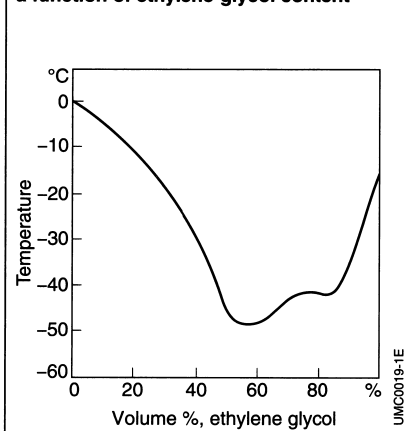
Radiator protection agents

The base material of radiator protection agents is for the most part 1,2-ethane-

Table 1: Requirements regarding water from the main based on the radiator manufacturer's recommendation

Property	Requirement
Appearance	colorless, clear
Deposit	0 mg/l
pH value	6.5 to 8.0
Total hardness	≤ 4.5 mmol/l
Chloride	≤ 100 mmol/l
Sulfate	≤ 100 mmol/l
Hydrogen carbonate	≤ 100 mmol/l

Figure 1: Freezing point of coolant as a function of ethylene-glycol content



diol (ethylene glycol), or less commonly 1,2-propandiol (propylene glycol). Both glycols have very good water-soluble properties. When mixed with water from the main, they lower the freezing point and raise the boiling point. The optimum level of these effects is achieved with a volume percentage of approximately 50 % (Figure 1). The boiling point is furthermore dependent on the operating pressure in the cooling system (Figure 2). The operating pressure of a passenger-car cooling circuit is in normal operation typically approximately 1.5 bar overpressure, under full load and in high-performance engines up to 2.5 bar and higher.

In the strict sense, with glycol-water mixtures there is not a defined freezing point, but an ice flaking point at which ice crystals begin to separate out in the fluid. At this temperature, the fluid medium can still be pumped through the cooling circuit.

The heat capacity of glycol-water mixtures is reduced compared with water from the main.

Since water/glycol mixtures demonstrate similar corrosivity to water from the main compared with metals, inhibitor systems are additionally admixed with the radiator protection agents.

Types of radiator protection agents

Radiator protection agents can essentially be classified into three groups, which differ with regard to the inhibitor systems used.

Conventional radiator protection agents containing silicate

Inhibitors and additives

These radiator protection agents contain primarily inorganic inhibitors such as, for example, silicate, nitrite, nitrate or molybdate. Organic inhibitors such as toluene or benzotriazole are also used.

To keep the pH value at the desired alkaline level over the entire period of use, buffer substances such as borate, phosphate, benzoate or imidazole are added to the radiator protection agent.

Radiator protection agents are completed by the addition of cleaning agents (e.g., sulfonates), antifoaming agents, and dyes.

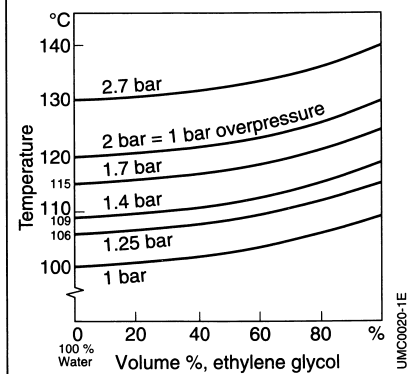
Conventional radiator protection agents containing silicate have proved themselves in practice over decades. Particularly the proportion of silicate has counteracted the dangerous hot corrosion of aluminum engine through the formation of thin, protective coatings.

Service life

However, the drawback of this type of radiator protection agent is that it is not possible to select the concentrations of the inhibitor systems at any level owing to the low solubility of some additives. On the other hand, the inorganic inhibitors are reduced during driving operation; however, the necessary minimum concentrations of inhibitors may not be undershot here. Consequently, the service life of such a radiator protection agent is limited to around three years driving or 100,000 km (for passenger cars). This service life can be significantly reduced for example by the high thermal load of the cooling circuit in modern high-performance engines.

The limited service life of conventional radiator protection agents does not match the wishes of the automobile and engine manufacturers to extend the service intervals of vehicles. The manufacturers of

Figure 2: Boiling-point curves of water/ethylene-glycol mixtures with different ethylene-glycol content as a function of pressure



radiator protection agents are therefore responding with the development of "OAT products".

OAT radiator protection agents

OAT radiator protection agents (Organic Acid Technology) contain a combination of organic inhibitors instead of conventional inorganic inhibitors. Typical combinations are aliphatic mono- and dicarboxylic acids, azelaic acid and aromatic carboxylic acids, sebacic acid and benzotriazole. The combination of an inorganic inhibitor and carboxylic acid is also used. Pure OAT products, however, contain no silicate.

In contrast to conventional radiator protection agents containing silicate, the inhibitors in OAT radiator protection agents decrease much more slowly in driving operation. This enables the service life of the radiator protection agents to be increased and the maintenance requirement to be reduced.

Hybrid radiator protection agents

To combine the good long-term durability of OAT radiator protection agents with the good hot-corrosion protection of silicate inhibitors, hybrid radiator protection agents have been developed which contain a harmonized combination of both inhibitor types. In this way, local hot corrosion can be prevented by the rapid formation of silicate protective coatings, while good long-term corrosion protection, even after the consumption of silicate inhibitors, is provided by the more stable OAT inhibitors.

This on the one hand satisfies the demand for low-maintenance coolants, but on the other hand also takes into account the high thermal loads of modern engine technologies. Therefore these hybrid radiator protection agents, sometimes also called Si-OAT radiator protection agents, are today frequently used by major vehicle manufacturers for the initial fill.

Brake fluids

Requirements

Passenger cars and light commercial vehicles are equipped with hydraulic brake systems. A hydraulic medium transmits and converts the braking force from the brake pedal in the brake master cylinder into hydraulic pressure, which is routed via brake lines to the wheel brake and there converted back into braking force. Vehicle-dynamics control systems such as the antilock braking system or the electronic stability program can additionally modulate the hydraulic pressures at each wheel to prevent the wheels from locking through pressure reductions or to stabilize the handling through additional pressure build-ups. This medium is therefore subject to the following requirements.

Low compressibility

The basic requirement with regard to the hydraulic medium brake fluid for effective power transmission is very low compressibility. The liquid state of aggregation must be maintained in the entire operating range for this purpose.

Water-absorption capacity

During driving water can penetrate into the brake system at the wheels, particularly through the flexible brake hoses usually made from rubber material. If this water is not bound by the brake fluid, com-

pressible water-vapor bubbles can form at temperatures above 100 °C. Consequently, no further pressure can be generated when the brake pedal is pressed because initially the compressible water vapor is compressed. This failure of the brakes is called “vapor lock”. Consequently, the market has seen the widespread acceptance of brake fluids which are completely mixable with water and exclude the presence of unbound, free water. Mineral-oil-based brake fluids also commonly used in the past have in practical terms been completely superseded by chemically glycol-based brake fluids (market share is approximately 99 %).

Standards for brake fluids

Brake fluids must comply with stringent requirements to ensure reliable brake-system operation. Their properties are defined in various standard requirements which are very similar in terms of content (SAE J1703 [1], SAE J1704 [2], FMVSS 116 [3], ISO 4925 [4], JIS K2233 [5]). Exclusively for the vehicle operating medium brake fluid the performance requirements of FMVSS 116 (Federal Motor Vehicle Safety Standard) became mandatory in the USA and often serve as a universal reference. In this standard the US Department of Transportation (DOT) – comparable with other standards – has defined specific ratings for salient properties (Table 1).

Table 1: Classification of brake fluids with different chemical bases

Chemical base Classification	Glycol ether				Silicone	Mineral oil
Suggested color	Colorless to amber				Purple	Green
FMVSS 116	DOT 3	DOT 4		DOT 5.1	DOT 5	
ISO 4925 class	3	4	6	5-1		ISO 7308
SAE	J 1703	J 1704 Standard	J 1704 Low viscosity		J 1705	
JIS K2233 class	3	4	6	5		
Dry boiling point [°C]	>205	>230	>250	>260	>260	>235
Wet boiling point [°C]	>140	>155	>165	>180	>180	
Viscosity at -40 °C [mm²/s]	<1500	<1800 (DOT) <1500 (SAE, ISO, JIS)	<750	<900	<900	<2000

Characteristics

Boiling point

During braking the temperature at the wheel brake cylinders increases, in extreme cases (e.g., frequent braking on long downhill gradients) even into ranges around 200 °C. Vapor bubbles can form at temperatures above the operating medium's instantaneous boiling point. The brake fluid transforms at least partly into a vapor state that is too compressible for reliable braking operation. It is then no longer possible to operate the brakes as is the case with the presence of water-vapor bubbles; "vapor lock" brake failure occurs.

Dry boiling point

In their new state (e.g. in sealed containers) brake fluids typically contain no or very little water, i.e. the water content is typically < 0.2 %. Fluids, particularly in steel containers, sometimes reach values of < 0.03 %. The equilibrium boiling point in this new state is called the dry boiling point (ERBP: Equilibrium Reflux Boiling Point).

The boiling point of the hygroscopic glycol-based brake fluid decreases significantly as a result of water absorption. Consequently, the risk of a "vapor lock" increases accordingly at high wheel-brake temperatures.

Depending on operation and the climate region, typically a water-content increase of approximately 1 to 1.5 % per year is assumed.

Wet boiling point

The equilibrium boiling point of the brake fluid after a precisely specified procedure of water absorption is defined as the wet boiling point (wERBP: wet Equilibrium Reflux Boiling Point) and is intended to depict the boiling point of a brake fluid used for two to three years.

The brake fluid to be tested is removed after roughly one day from a very humid climatic chamber to measure the boiling point when a reference brake fluid has reached a 3.7 % water content. The boiling point of the operating medium continues to drop if further water is absorbed. Thus the brake-failure risk increases at high wheel-brake temperatures. This is

the main reason why the brake fluid typically should be replaced every two years. Figure 1 shows by way of example the drops in boiling point that result from water absorption in two brake fluids.

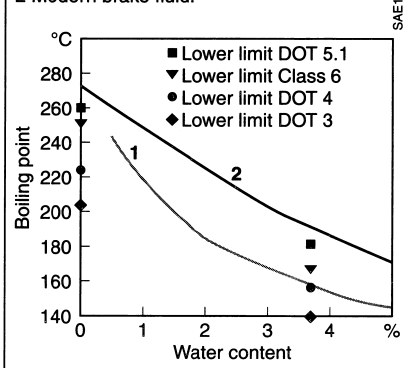
A significant development trend on the part of the brake-fluid manufacturers in previous years was to increase the boiling point with increased water content.

Viscosity

The viscosity describes the resistance to flow of the operating medium. As the viscosity increases, so too the pressure drops increase as the medium flows through lines and throttle points, e.g., in valves and pumps of the vehicle-dynamics control systems. This reduces, for example, the dynamics of the pressure build-up, which can reduce the effectiveness of the electronic stability program. Therefore, generally speaking, low cold viscosities and low dependence of viscosity on temperature are desirable. This is specified by upper limits for the viscosity at very low temperatures at -40 °C (cold viscosity). It should be as low as possible in order to facilitate high dynamics on the part of the pressure build-up. Reducing the viscosity at very low temperatures was a further development trend on the part of the brake-fluid manufacturers in past years. Figure 2 shows by way of

Figure 1: Reduction of the boiling point with increasing water content from the example of two brake fluids

1 SAE reference brake fluid [6], [10],
2 Modern brake fluid.



example the dependence of viscosity on temperature for two brake fluids. This has a clearly non-linear characteristic particularly at low temperature. As a general "rule of thumb": a 1 K reduction in temperature produces approximately a 10 % increase in viscosity. Likewise, approximately a 10 % increase in viscosity results at low temperature additionally from a 1 % increase in water content. This effect is another reason why the brake fluid should be changed.

Corrosion protection

FMVSS 116 (and also all the other standards) stipulates that brake fluids shall exercise no corrosive effect on those metals generally employed in brake systems. Appropriate additives guarantee the required corrosion protection, which is checked beyond required upper limits in high-temperature storage tests of metal plates in dry and in wet brake fluids.

Elastomer swelling

Elastomers are used as seals and gaskets in the brake system, for example in the brake master cylinder, in vehicle-dynamics control systems, and in hoses. Although a small amount of elastomer swelling is desirable and taken into account in the design of the sealing rings, it should not exceed 10 % under any circumstances; otherwise, it has a negative effect on the strength of these components.

Typically, EPDM is used as elastomer material for glycol-based (and also for rarely used silicone-oil-based) brake fluids. Mineral-oil-based brake fluids are not compatible with EPDM and required different elastomer materials (e.g., FKM).

Even minute levels of mineral-oil contamination (e.g., hydraulic fluids, mineral-oil-based brake fluid) of glycol-based brake fluid can destroy EPDM elastomers (e.g., sealing elements). This can ultimately lead to brake failure. To avoid such contamination, high standards of cleanliness must be maintained for the brake fluids used during service filling [10].

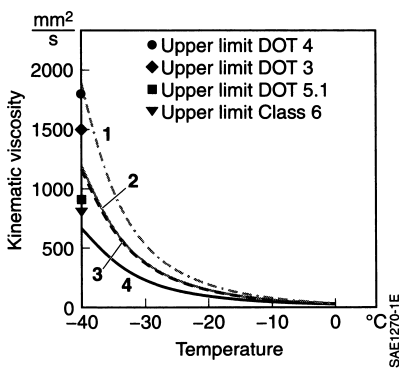
Chemical composition

Glycol-ether fluids

Most brake fluids are based on glycol-ether compounds. These generally consist of monoethers of low polyethylene glycols. The chemical formulation of the ISO and at the same time the SAE reference brake fluids is specified in detail in ISO 4926 [6]. Glycol components can be used to produce a brake fluid which conforms to DOT 3 requirements (Table 1), the downside is that their boiling point drops relatively quickly with water absorption. DOT 4 and DOT 5.1 brake fluids are characterized by the fact that glycol-ether borates are also used. This enables the water to be chemically bound. In this way, the boiling point drops much more slowly compared with DOT 3 fluids. The wet boiling point rises significantly (Figure 1). Thus, the risk of a "vapor lock" decreases significantly when the wheel brake is subject to increased water content and high temperatures.

Figure 2: Dependence of viscosity on temperature and water content from the example of two brake fluids

- 1 SAE reference brake fluid with 5 % water,
- 2 SAE reference brake fluid without water [6], [10],
- 3 Modern brake fluid with 5 % water,
- 4 Modern brake fluid without water.



Future development trends

ISO 4925 defines a further quality class, namely "Class 6". This quality is better than that of DOT 4 and is characterized by its particularly low cold viscosity (Table 1). It constitutes the first combination of the two development trends of increased wet boiling point and reduced cold viscosity.

This trend towards higher wet boiling points and low cold viscosities will continue and combine. Increasingly "modern" brake fluids are being developed and offered.

In addition, the functionality of the vehicle-dynamics control systems – e.g., vehicle-speed controller including brake intervention as ACC (Adaptive Cruise Control) and AEB (Automated Emergency Braking) – is constantly increasing. The road towards (semi-)automated driving is clear. The frequency of use of the vehicle-dynamics control systems thereby increases. While a pump element in the ABS (antilock braking system) was only rarely active, its duration of use and hence also its operating load increase significantly by way of TCS (traction control system) and electronic stability program with comprehensive additional functions right up to automated driving.

This also renders necessary in future in the brake-fluid standards requirements, e.g., with regard to the lubricity of the brake fluid, in order to reduce wear. The same also applies to avoid noise when operating the brake master cylinder and clutch, which in the meantime is increasingly operated in manual transmissions with brake fluid. Working groups, particularly within the SAE and ISO, and the pfp TriNoWe (public funded project Tribology in Standards Worldwide; German: Tribologie in Normen Weltweit), are working on further developments of the standards.

Other types of brake fluids

Silicone-oil fluids (SAE J1705, [7])

Silicone-oil-based brake fluids are now only used in a few special applications, e.g. in motor sport, in which very short maintenance intervals are possible.

The disadvantage of these products is – aside from the high price – the higher compressibility. Furthermore, they are – like mineral oils – not hygroscopic and therefore not able to absorb water that

penetrates into the brake system. For this reason, they are also called "Low Water Tolerant" brake fluids.

Mineral-oil fluids (ISO 7308, [8])

Additives for improving the fluid properties (e.g., lubrication) are known from mineral-oil technology. These were used previously for example in vehicles with central hydraulics, but are now no longer significant because they – like silicone-based fluids – cannot bind water.

Mineral-oil fluids should never be added to brake systems which are designed for glycol fluids (or vice versa), as this would destroy the elastomer seals. For this reason, the FMVSS suggests a color-coding system for the various brake fluids (see Table 1).

References

- [1] SAE J1703: Motor Vehicle Brake Fluid.
- [2] SAE J1704: Motor Vehicle Brake Fluid Based Upon Glycols, Glycol Ethers and the Corresponding Borates.
- [3] FMVSS 116: Federal Motor Vehicle Standard No. 116: Motor Vehicle Brake Fluids.
- [4] ISO 4925: Road vehicles – Specification of non-petroleum-base brake fluids for hydraulic systems.
- [5] JIS K2233: Non-petroleum base motor vehicle brake fluids.
- [6] ISO 4926: Road vehicles – Hydraulic braking systems – glycol-based reference brake fluids.
- [7] SAE J1705: Low Water Tolerant Brake Fluids.
- [8] ISO 7308: Road vehicles – Petroleum-based brake fluid for stored-energy hydraulic brakes.
- [9] SAE J1709: (Historical) European Brake Fluid Technology.
- [10] SAE J1706: Production, Handling and Dispensing of SAE J1703 Motor Vehicle Brake Fluids and J1704 Borate Ester Based Brake Fluids.

Fuels

Overview

Composition

Fuels consist of a multitude of hydrocarbons, primarily of various paraffins and aromatic compounds.

Paraffins

Paraffins are linear, isoparaffins branched and cycloparaffins closed-chain hydrocarbons which have no multiple bonds, and are therefore chemically less reactive.

Aromatic compounds

Aromatic compounds are characterized by the property that the molecules have at least one ring, in which electrons are distributed around the entire ring. Benzene is a typical example. Because the electrons are not located as a double bond between two carbon atoms, aromatic compounds are likewise less active.

Production

Gasoline and diesel fuel are produced from crude oil. First the crude oil is desalted and distilled, where the hydrocarbons are separated, depending on the molecule size, into different fractions.

Distillation under atmospheric pressure

The process is begun with a distillation of the crude oil under atmospheric pressure, in which different substance streams of predominantly volatile components are produced. Depending on their composition, these substance streams form the basis for liquefied petroleum gas (propane and butane), gasoline, kerosene, and for the more volatile part of diesel fuel and fuel oil.

Distillation under vacuum

The remaining residue from the atmospheric distillation is further separated under vacuum. These fractions form the basis for the heavier components in diesel fuel and fuel oil and for light and heavy marine fuels. During vacuum distillation too a residue is left, which is processed into heavy fuel oil and bitumen.

Further processing

The ratio and the quality of the product fractions obtained by distillation do not yet meet market requirements such that further processing in complex chemical processes is required in the refinery.

Aside from the widely known desulfurization, larger molecules can be cracked purely thermally or in the presence of catalysts (thermal cracking, catalytic cracking). If this process is performed in the presence of hydrogen, it is known as hydrocracking. Additional components can be produced from heavy oil for diesel fuel by hydrocracking.

The structure of molecules is modified in other refinery processes. During reforming aromatic compounds are produced from saturated aliphatic molecules, where hydrogen that is needed for hydrocracking and desulfurization is created. Linear molecules can be converted into branched molecules by isomerization processes. These increase the octane number of gasoline and the low-temperature resistance of diesel fuel.

Addition of biofuels and additives

Today proportions of biofuels are added to many fossil fuels worldwide. Certain properties of biofuels can be specifically used to satisfy the fuel-grade requirements. For example, bioethanol increases the octane number of gasoline and adding biodiesel to diesel fuel ensures sufficient lubricity. Additives are often also used to further improve product quality. Additives are active substances which are added to specifically adjust certain fuel properties.

Characteristics

Boiling range and ignition temperature

The boiling range of gasoline is between 30 °C and 210 °C, and that of diesel fuel between 160 °C and 370 °C. Diesel fuel ignites on average at approximately 350 °C (lower limit 220 °C), which is very early in comparison with gasoline (on average 500 °C).

Calorific value

Normally the net calorific value H_n (former term: lower calorific value) is specified for the energy content of fuels; it corresponds to the usable heat quantity released during full combustion.

The gross calorific value H_g (former term: upper calorific value), on the other hand, specifies the total reaction heat released and therefore comprises, as well as the mechanically usable heat, the latent heat created in the water vapor. However, this component is not used in the vehicle.

At 42.9 – 43.1 MJ/kg, the net calorific value of diesel fuel is slightly higher than that of gasoline (40.1 – 41.9 MJ/kg).

Fuels or fuel constituents containing oxygen (oxygenates), such as alcohol fuels, ether, or fatty acid methyl ester, have a lower calorific value than pure hydrocarbons because the oxygen bonded in them does not contribute to the release of energy in the combustion process. Engine performance comparable with conventional fuels can therefore only be achieved with fuels containing oxygen by increased fuel consumption.

Calorific value of air/fuel mixture

This is the calorific value of the combustible air/fuel mixture. It is dependent on the air/fuel ratio and determines the engine's power output. With a stoichiometric air/fuel ratio, this is roughly 3.5 - 3.7 MJ/m³ for all liquid fuels and liquefied petroleum gases.

Sulfur content

In the interests of reducing sulfur-dioxide emissions (SO₂) and protecting the catalytic converters for exhaust-gas treatment, the sulfur content of gasolines and diesel fuels was limited for vehicle applications on a Europe-wide basis to 10 mg/kg as from 2009. Fuels which adhere to this limit value are known as "sulfur-free fuels". In this way, the final stage of the desulfurization of fuels is achieved. Prior to 2009 only low-sulfur fuel (sulfur content < 50 mg/kg) – introduced at the beginning of 2005 – was permitted for use in Europe. Germany has led the way in the desulfurization of fuels and already back in 2003 had established sulfur-free fuel by way of tax measures.

In the USA, the limit value for the sulfur content of gasolines commercially available to the end user has since 2006 been set at max. 80 mg/kg, although an upper average value of 30 mg/kg for the total amount of sold and imported fuel is in place. Individual states, California for example, have laid down lower limits.

Furthermore, 2006 saw the beginning of the introduction in the USA of sulfur-free diesel fuels for highway applications (ULSD, Ultra Low Sulfur Diesel, sulfur content max. 15 mg/kg). By the end of 2009, however, only 20 % of fuels would have a sulfur content of max. 500 mg/kg.

The sulfur content of certification fuels is subject to separate regulations.

Gasoline (petrol)

Fuel grades

Super fuels in Germany

In Germany two Super fuels with 95 octane are available; these differ in their ethanol content and may contain a maximum 5 % (V/V, % by volume) ethanol (Super) or 10 % (V/V) ethanol (Super E 10). A Super Plus fuel with 98 octane is also available. Individual providers markets fuels with an octane number of at least 100 (V-Power Racing 100, Super 100plus, Ultimate 102), which in comparison with the Super fuels with 95 octane are changed in terms of basic quality and additives.

Fuels in the USA

Three different types of fuel are sold in the USA: Regular (92 octane), Premium (94 octane), and Premium Plus (98 octane). Fuels sold in the USA usually contain 10 % (V/V) ethanol. The addition of components containing oxygen increases the octane number and satisfies the requirements of modern, higher-compression engines with regard to improved knock resistance.

Reformulated gasoline is the term used to describe gasoline which, through its altered composition, generates fewer evaporative and pollutant emissions than conventional gasoline. The demands placed on reformulated gasoline are laid down in the USA in federal law (Clean Air Act). In comparison with standard quality, this legislation prescribes, for example, lower limits for vapor pressure, benzene and overall aromatic content, and final boiling point. It also prescribes the use of additives to keep the intake system free of contamination and deposits (detergents).

Fuel standards

European Standard EN 228 [1] defines the requirements for unleaded fuel for use in spark-ignition (SI) engines. Further country-specific characteristics are set out in the national appendices to this standard (Table 1). Leaded gasolines are no longer permitted in virtually all markets worldwide and by 2014 were only available in a few countries, e.g., Algeria.

The US specifications defining fuels for gasoline engines are contained in

ASTM D4814 (American Society for Testing and Materials) [2].

Most fuels for gasoline engines which are sold today contain components which in turn contain oxygen (oxygenates). In this respect, ethanol has gained in importance in Europe as well, since the "EU Bio-fuels Directive" stipulates minimum content figures for renewable fuels. But also those ethers which can be produced from methanol or ethanol MTBE (methyl tertiary butyl ether) and ETBE (ethyl tertiary butyl ether) are used, of which, depending on the quality of the basic fuel, up to 22 % (V/V) may be added in Europe. A mixture of alcohols and ethers is also permitted. In Europe the overall quantity of oxygenates is however limited by the stipulation of a maximum permitted oxygen content. Thus, E 10 in Europe may exhibit a maximum oxygen content of 3.7 % (m/m, mass percent), which is already exhausted with the addition of 10 % ethanol. When ether is added, the maximum volume percentage of ethanol must be reduced accordingly. Worldwide many countries have defined minimum quotas for biogenic components in gasolines, which are achieved for the most part with bioethanol.

Adding alcohols increases the volatility at higher temperatures. Furthermore, alcohol can damage materials in the fuel system, lead for example to elastomer swelling, and trigger alcoholate corrosion on aluminum parts. Depending on the alcohol content and the temperature, demixing may result even in response to only small amounts of water. During phase separation a new aqueous phase is formed with part of the alcohol from the fuel. The use of fuels containing alcohol therefore require additional measures on the vehicle and in the filling-station network.

The ethers do not encounter the problem of demixing. The ethers, which have a lower vapor pressure, a higher calorific value and a higher octane number than ethanol, are chemically stable components with good material compatibility. They therefore demonstrate advantages, both from an engine and a logistical standpoint. For reasons of sustainability, because of the CO₂ savings to be attained and due to the laying down of quotas for biogenic fuels, ETBE is used in the main.

Where the ethanol content had long been limited in the European gasoline standard EN 228 to 5 % (V/V) (E5), the 2013 edition contains in the first instance a specification for ethanol at 10 % (V/V) (E10). At present in the European market not all vehicles are yet equipped with materials to allow operation with E10. As the second quality, therefore, an exemption grade with a max. ethanol content of 5 % (V/V) is retained. In Germany the new edition DIN EN 228:2014 replaces the draft standard E DIN 51626-1:2010-04 [3], which was published in April 2010, to enable E10 to be introduced early to the German market. E10 is conceived as the main grade in the gasoline market, but as yet has failed to find favor with consumers. E10 is up to now only available in Europe in Germany, France, and Finland (as at the start of 2016).

In the USA most gasolines contain ethanol at 10 % (V/V) (E10), for which a vapor pressure higher by approximately 7 kPa is permitted according to the American standard ASTM D4814. A vapor pressure that is up to 8 kPa higher is also permitted in EN 228 (2013), depending on the ethanol content.

Characteristics

Density

European standard EN 228 limits the density of gasolines to 720 - 775 kg/m³.

Octane number

The octane number defines the gasoline's antiknock quality (resistance to pre-ignition). The higher the octane number, the greater the resistance to engine knock. Iso-octane (trimethyl pentane), which is extremely knock-resistant, is assigned the octane number 100, while n-heptane, which is extremely knock-susceptible, is assigned the number 0.

The mixture of iso-octane and n-heptane, which exhibits the same knock resistance as the fuel to be tested, is determined in a standardized test engine. The proportion in % by volume of iso-octane in the mixture of iso-octane and n-heptane corresponds to the fuel's octane number.

Research and motor octane numbers

The octane number determined in testing using the Research Method ([4]) is the RON (research octane number). It serves as the essential index of acceleration knock. The octane number determined in testing using the Motor Method ([5]) is the MON (motor octane number). The MON basically provides an indication of the tendency to knock at high speeds.

The Motor Method differs from the Research Method by using preheated mixtures, higher engine speeds and variable ignition timing. This places more stringent thermal demands on the fuel under examination. MON figures are lower than those for RON.

Increasing the antiknock quality

Normal (untreated) straight-run gasoline displays a low antiknock quality. Only by mixing such gasoline with different knock-resistant refinery components (reformed components, isomerisates) is it possible to produce fuels with high octane numbers suitable for modern, high-compression engines. Even components containing oxygen such as alcohols and ether increase the antiknock quality. They are therefore specifically added for this purpose.

Additives containing metal to increase the octane number (e.g. MMT: methylcyclopentadienyl manganese tricarbonyl) form ashes during combustion. MMT is therefore excluded in EN 228 (2013) by a limit value for manganese in the trace range.

Characteristics of fuel volatility

The volatility of gasoline has both upper and lower limits. The lower limitation is chosen so that the fuel contains sufficiently volatile components to ensure a reliable cold start. However, the fuel's volatility may not be so high that at higher temperatures the fuel supply is interrupted by the formation of gas bubbles (vapor lock) and as a result difficulties when hot starting or problems when driving are encountered. Yet another factor is environmental protection, which demands that evaporative losses be kept low.

Fuel volatility is defined by different characteristic quantities. EN 228 defines

for E 5 and E 10 ten different volatility classes distinguished by various levels of boiling curve, vapor pressure and VLI (vapor-lock index). To meet special requirements stemming from variations in climatic conditions, European countries can incorporate specific individual classes into their own national appendices in the standard. Different values are laid down for summer and winter.

A similar provision is included in the American gasoline standard ASTM D4814. Comprehensive tables describe the scope of application of the volatility classes, depending on the calendar month and state.

Boiling curve

In order to assess the fuel in vehicle operation, it is necessary to view the individual areas of the boiling curve separately. EN 228 therefore contains limit values laid down for the volume of fuel that vaporizes at 70 °C, at 100 °C and at 150 °C. The volume of fuel that vaporizes at up to 70 °C must achieve a minimum volume in order to ensure that the engine starts easily when cold (this was previously important above all for vehicles with carburetors). However, the volume of fuel that vaporizes must not be too great either, otherwise vapor bubbles may be formed when the engine is hot. While the volume

of fuel that vaporizes at up to 100 °C determines the engine's warm-up characteristics, this factor's most pronounced effects are reflected in the acceleration and response provided by the engine once it warms to normal operating temperature. The volume of fuel that vaporizes at up to 150 °C should be high enough to minimize dilution of the engine oil. In particular when the engine is cold, the non-volatile gasoline components find it difficult to vaporize and can pass from the combustion chamber via the cylinder walls into the engine oil.

Vapor pressure

Fuel vapor pressure as measured at 37.8 °C (100 °F) in accordance with EN 13016-1 [6] is primarily a characteristic quantity with which the safety requirements in the vehicle's tank are established. Vapor pressure has upper and lower limits in all specifications. In Germany, for example, it is max. 60 kPa in summer and max. 90 kPa in winter.

In order to configure a fuel-injection system, it is also important to know the vapor pressure at higher temperatures (80...120 °C) because a rise in the vapor pressure due to the admixture of alcohol becomes apparent only at higher temperatures. If the fuel vapor pressure rises

Table 1: DIN EN 228 (August 2017): Selected requirements of gasolines

Requirements	Unit	Characteristic	
Quality		Super E10	Super
Knock resistance: RON/MON for Super, min.	–	95/85 ¹	
Density (at 15 °C), min./max.	kg/m ³	720/775	
Sulfur, max.	mg/kg	10	
Oxygen content, max.	% (m/m)	3.7	2.7
Ethanol content, max.	% (V/V)	10.0	5.0
Benzene, max.	% (V/V)	1	
Lead, max.	mg/l	5	
Volatility			
Vapor pressure, summer, min./max.	kPa	45/60	
Vapor pressure, winter, min./max. ²	kPa	60/90	
Evaporated volume at 70 °C, summer, min./max.	% (V/V)	22/50	20/48
Evaporated volume at 70 °C, winter, min./max.	% (V/V)	24/52	22/50
Evaporated volume at 100 °C, min./max.	% (V/V)	46/72	46/71
Evaporated volume at 150 °C, min./max.	% (V/V)	75/–	
Final boiling point, max.	°C	210	
VLI ³ transitional period ⁴ , max. ²		1 164	1 150

¹ Applicable to Super Plus: RON/MON, min. 98/88; ethanol content, max. 10.0 % (V/V).

² National values for Germany. ³ VLI Vapor-lock index. ⁴ Spring and fall.

above the fuel-injection system pressure, for example during vehicle operation due to the effect of the engine temperature, this may result in malfunctions caused by vapor-bubble formation.

Vapor/liquid ratio

The vapor/liquid ratio is a measure of a fuel's tendency to form bubbles. It refers to the volume of vapor generated by a specific quantity of fuel at a defined back pressure and a defined temperature. A drop in pressure (e.g. when driving over a mountain pass), or an increase in temperature, will raise the vapor/liquid ratio and with it the possibility of operating problems. ASTM D4814 lays down, for example for each volatility class, a temperature at which a vapor/liquid ratio of 20 must not be exceeded.

Vapor-lock index

The vapor-lock index (VLI) is the mathematically calculated sum total of ten times the vapor pressure (in kPa at 37.8 °C) and seven times the volume percentage of fuel that vaporizes at up to 70 °C. It is possible with this additional limit value to restrict the volatility of the fuel further with the result that the two maximum values for vapor pressure and boiling data cannot be achieved in the course of its production.

Additives

Additives can be added to improve fuel quality in order to counteract deteriorations in engine performance and in the exhaust-gas composition during vehicle operation. The packages generally used combine individual components with various attributes.

Care and precision are required both when testing additives and in determining their optimal compositions and concentrations. Undesirable side-effects must be avoided. Basic additives are added at the refinery to protect the plants and to ensure a minimum fuel quality. To further improve quality, brand-specific multifunction additives can be added at the refinery's filling stations when the road tankers are filled (end-point dosing).

Detergents

The setting and preparation of the mixture composition of air and fuel in the combustion chamber is pivotal for trouble-free driving and minimization of pollutants in the exhaust gas. Keeping the entire intake system (fuel injectors, intake valves) clean is an important prerequisite for meeting the optimized conditions in the new state. Effective cleaning additives (detergents) help to keep surfaces clean and remove existing coatings.

Corrosion inhibitors

The ingress of water/moisture may lead to corrosion in fuel-system components. Corrosion can be effectively eliminated by the addition of corrosion inhibitors, which form a thin protective film on the metal surface.

Aging stabilizers

Oxidation inhibitors (antioxidants) are added to fuels to improve their stability during storage. They prevent rapid oxidation caused by oxygen in the air.

Blend components for fossil gasolines

Bioethanol

Production from sugar and starch
Bioethanol can be obtained from products containing sugar and starch and is the most widely produced biofuel worldwide. Plants containing sugar (sugar cane, sugar beet) are fermented with yeast, the sugar fermenting to form ethanol in the process.

When bioethanol is obtained from starch grains such as corn, wheat or rye are pretreated with enzymes in order to partially split the long-chain starch molecules. A split into dextrose molecules occurs during the subsequent likewise enzymatically performed saccharification. Bioethanol is created in a further process step by means of fermentation with yeast.

Production of lignocellulose

Enzymes can also be used to produce bioethanol from lignocellulose. Lignocellulose forms the structural framework of the plant cell wall and contains the chief constituents lignin, hemicellulose and cellulose. The advantage of this process is that the entire plant can be used and not

just that part containing sugar or starch. Because of this new approach, this product is also referred to as 2nd-generation bioethanol. However, the enzymes can split cellulose, but not lignin. The enzymes developed to date are still highly sensitive with regard to the usable biomass such that this process is currently still of no economic significance.

Use of bioethanol

Bioethanol is, on the basis of its properties, highly suitable for admixing in gasolines, particularly in order to increase the octane number of pure gasoline. That is why virtually all gasoline standards permit the addition of ethanol as a blend component. Even the biofuel policy of the European Union leads one to expect that the market penetration and the proportion of bioethanol in gasolines will continue to rise if it is guaranteed that bioethanol is created on a sustained basis and not in competition with foods.

Bioethanol can also be used as a pure fuel in spark-ignition engines in flexible fuel vehicles (FFVs). These vehicles can run both on gasoline and on pure ethanol as well as any mixture of gasoline and ethanol. A gasoline with 85 % (V/V) ethanol is offered in some European countries, mainly Germany, France, and Sweden. On account of the cold-start problems at low temperatures, the maximum ethanol concentration is sometimes reduced depending on the local climate. Fuel samples exhibit typical ethanol contents of 75 to 85 % (E85) in summer and 65 to 85 % in winter. The quality of E85 is defined for Europe in the fuel specification CEN/TS 15293 [7] and in the USA in ASTM D5798 [8]. In Brazil gasolines are basically only sold as ethanol fuels, which are also called gasohol and contain 22 to 26 % (V/V) ethanol. Unblended ethanol, E100, is also available at many Brazilian filling stations, but this contains 7 % (m/m) water.

Methanol

Production

Methanol is primarily produced not by regenerative means from fossil energy sources such as coal and natural gas, and therefore makes no contribution to reducing CO₂ emissions. Countries such as China, who plan to cover their high fuel demand in part from coal, will increasingly be opting for methanol in the future. Here, M15 appears to represent an upper limit for use in conventional spark-ignition engines. In China, an M85 analogous to E85 is being discussed for flexible fuel vehicles.

Use of methanol

With the same alcohol content, methanol fuels have a significantly greater corrosive effect on ferrous alloys than ethanol fuels. Even demixing occurs much more quickly when the fuel is fouled by water. Because of the negative experiences of methanol fuels in Germany during the oil crisis in 1973 and also due to its toxicity, the use of methanol as a blend component has again been abandoned in Germany. From a worldwide point of view, only occasional methanol admixtures are being conducted at present, and then for the most part with a content of well below 5 % (M5).

Diesel fuel

Fuel standards

The standard that lays down the requirements for diesel fuels in Europe is EN 590 [9]. The most important characteristics are set out in Table 2. Even the premium qualities additionally sold at some filling stations (e.g. Super diesel, Ultimate, V-Power diesel) satisfy this standard. There are differences in the basic quality, in the addition of paraffinic fuel components, and in the additives.

Up to 7 % (V/V) biodiesel (FAME) can be admixed with diesel fuel in accordance with EN 590, where the quality of the biodiesel is stipulated by standard EN 14214 [10]. The admixture of biodiesel improves lubricity, but also reduces oxidation stability. For the purpose of safeguarding sufficient oxidation stability, EN 590 was supplemented in 2009 to include the parameter induction time, which is a measure of the fuel's aging reserve and must satisfy a minimum requirement.

The US standard for diesel fuels, ASTM D975 [11], specifies a smaller number of quality criteria and defines less stringent limits. It permits the admixture of max. 5 % (V/V) biodiesel, which must satisfy the requirements of standard ASTM D6751 [12].

Characteristics

Cetane number

The cetane number (CN) expresses the ignition quality of the diesel fuel [13]. The higher the cetane number, the greater the fuel's tendency to ignite. As the diesel engine dispenses with an externally supplied ignition spark, the fuel must ignite spontaneously (auto-ignition) and with minimum delay (i.e. ignition lag) when injected into the hot, compressed air in the combustion chamber. Cetane number 100 is assigned to n-hexadecane (cetane), which ignites very easily, while slow-igniting α -methyl naphthalene is allocated cetane number 0. The cetane number of a diesel fuel is determined in a standardized CFR single-cylinder test engine with a variable compressor piston (CFR, Cooperative Fuel Research). The compression ratio of the fuel to be tested is measured at constant ignition lag. Reference fuels composed of cetane and heptamethylnonane, a component with the cetane number 15, are operated with the determined compression ratio. The proportion of cetane in the mixture is altered until the same ignition lag is obtained. The cetane proportion in percent specifies the cetane number.

Cetane numbers over 50 are desirable for optimum operation of modern engines, particularly under cold-starting conditions. High-quality diesel fuels contain a high proportion of paraffins with high cetane numbers. Conversely, aromatic compounds exhibit a low ignition quality.

Table 2: DIN EN 590 (October 2017): Selected requirements of diesel fuels

Requirements	Unit	Characteristic
Cetane number, min.	–	51
Cetane index, min.	–	46
Density (at 15 °C), min./max.	kg/m ³	820/845
Viscosity (at 40 °C), min./max.	mm ² /s	2.0/4.5
Sulfur content, max.	mg/kg	10
Lubricity, "wear scar diameter", max.	µm	460
FAME content, max.	% (V/V)	7
Oxidation stability	Insoluble ³ , max.	g/m ³
	Induction period (at 110 °C), min.	hours
Water content, max.	mg/kg	200
Total contamination, max.	mg/kg	24
CFPP ¹ in six seasonal classes, max. ²	°C	+5 to –20
Flash point	°C	> 55

¹ Filterability limit. ² Determined nationally, for Germany 0 to –20 °C.

³ Sum total of insoluble substances after aging.

Cetane index

Yet another parameter of ignition quality is provided by the cetane index, which is calculated on the basis of fuel density and various points on the boiling curve. This purely mathematical parameter does not take into account the influence of cetane improvers on ignition quality. In order to limit the adjustment of the cetane number by means of cetane improvers, both the cetane number and the cetane index have been included in the list of requirements in the standard EN 590. Fuels whose cetane number was increased with ignition improvers can exhibit a higher fuel consumption than fuels with the same high natural cetane number.

Boiling range

The boiling range of a fuel, i.e. the temperature range at which the fuel vaporizes, depends on its composition. A low initial boiling point is a prerequisite for a fuel suitable for use in cold weather, but also increases the risk of cavitation damage and results in greater system wear due to poorer lubrication properties.

If, however, the end of boiling is at the high end of the temperature scale, this can result in higher soot emissions and nozzle coking. This refers to the formation of deposits as a result of the chemical decomposition of non-volatile fuel components in the spray hole and on the nozzle cone and the addition of combustion residues. The processes of carbon depositing are highly complex. Above all non-volatile fuel components from the refinery cracking processes contribute to coking.

The degree of coking, which is calculated from the measured hydraulic throughflow of the uncleaned and cleaned nozzle, is for a specified nozzle geometry a measure of the tendency of a fuel to form deposits in the injection nozzles. Conversely, it is possible to assess with a reference fuel how the design of different nozzle configurations influences the formation of carbon deposits.

When the end of boiling is higher, the ingress of fuel over the cylinder walls into the engine oil is greater. For this reason, the percentage of non-volatile fuel components should not be too high. Limiting the admixture of biodiesel to max. 7 % (V/V) in EN 590 has its roots in the high boiling point of biodiesel (320...360 °C).

Filterability limit

Precipitation of paraffin crystals at low temperatures can result in fuel-filter blockage, ultimately leading to an interruption of the fuel flow. In worst-case scenarios, paraffin particulates can start to form already at temperatures of 0 °C or even higher. The cold resistance of a fuel is assessed by means of the temperature at which the fuel filter becomes clogged under defined test conditions (CFPP, Cold Filter Plugging Point, filterability limit). Standard EN 590 defines different limit values for the CFPP in various classes, and can be defined by individual member states depending on the prevailing geographical and climatic conditions.

Formerly, owners sometimes added regular gasoline to their vehicle fuel tanks to improve the cold response of diesel fuel. This practice is no longer necessary now that fuels conform to standards, and this may in any case cause damage particularly in today's widely used systems with high-pressure fuel injection.

Flash point

The flash point is defined as the temperature at which the quantities of vapor which a combustible fluid emits to the atmosphere are sufficient to allow a spark to ignite the air/vapor mixture above the fluid. For safety reasons (e.g. for transportation and storage), diesel fuel must comply with Hazard Class A III, i.e. its flash point must be above 55 °C. Less than 3 % gasoline in the diesel fuel is sufficient to lower the flash point to such an extent that ignition becomes possible at room temperature.

Density

The energy content per unit of volume increases in the case of purely fossil diesel fuel produced only from crude oil with density. Assuming constant activation of the injectors (i.e. constant injected fuel quantity), the use of fuels with widely different densities causes variations in mixture ratios (change in the air/fuel ratio λ) due to fluctuations in calorific value. When an engine runs on fuel that has a high type-dependent density, engine performance and soot emissions increase; at low fuel density, these parameters drop. For this reason, the type-dependent density spread for diesel fuel is tightly limited.

When biodiesel is admixed with diesel fuel, the density of the mixture increases due to the higher density of biodiesel, but the energy content decreases due to the amount of oxygen contained.

Viscosity

Viscosity is a measure of a fuel's resistance to flow due to internal friction. Excessively low viscosity of the diesel fuel leads to increased return quantities in the injector, as a result of which the fuel-injection system is additionally heated up and in some cases the injected fuel quantity required for full load is no longer achieved. In addition, the risk of wear and cavitation erosion increases.

Much higher viscosity—for instance when pure biodiesel (FAME) is used—causes a higher peak injection pressure at high temperatures in non-pressure-regulated systems (e.g., unit injector systems) compared with petroleum diesel. The fuel-injection system may not therefore be applied with petroleum diesel to the permissible peak pressure. High viscosity also changes the spray pattern due to the formation of larger droplets. The droplet distribution in the combustion chamber is called the spray pattern. Since the fuel viscosity increases sharply in cold conditions, restrictions in fuel delivery may be experienced as a result of the reduction of the volumetric flow.

Lubricity

The hydrodynamic lubricity of diesel fuels is not as important as the lubricity in the mixed-friction range. The introduction of environment-compatible, hydrogenation-desulfurized fuels resulted in huge wear problems with distributor injection pumps in the field. Desulfurization also sees the removal of fuel components which are important to lubricity. Lubricity enhancers have to be added to many fuels to avoid these problems. The standard EN 590 prescribes a minimum lubricity, measured as wear of max. 460 μm in an oscillation wear test (High-Frequency Reciprocating Rig, HFRR [14]).

The minimum-lubricity requirement is also achieved when sufficient biodiesel, usually at least 2 % (V/V), is admixed with the diesel fuel.

Oxidation stability

The constituents of diesel fuels are subject to a more pronounced extent than gasoline to oxidation caused by oxygen in the air. This is the case particularly when biodiesel is admixed. Polymers and acids are produced during oxidation, which is also referred to as fuel aging. Aging polymers can directly give rise to the formation of coatings and to gumming in the fuel-injection system. Aging acids frequently react with metal ions to produce organic salts and soaps (carboxylates), which are insoluble in diesel fuel and therefore also form coatings.

It is important that the fuel, while it is in the vehicle, has a sufficient aging reserve. The aging reserve is measured as an induction period under the test conditions defined in EN 15751 [15] at 110 °C. When diesel fuel at the filling-station pump exhibits an aging reserve of at least 20 hours according to this test standard, it is ensured that the fuel is stable enough for normal use. In the past, failures caused by fuel aging have only occurred in isolated cases when vehicles were subject to very long idle periods, in special agricultural applications, or also in the construction sector.

Total contamination

Total contamination refers to the sum total of undissolved foreign particulates in the fuel, such as sand, rust, and undissolved organic components, including aging polymers of the fuel. The standard EN 590 permits a maximum total contamination of 24 mg/kg. Very hard silicates which occur in mineral dust are specially damaging to high-pressure fuel-injection systems with narrow gap widths. Even a fraction of the permissible overall contamination level of hard particulates would cause erosive and abrasive wear (e.g. at the seats of solenoid valves). Wear of this nature results in valve leakage, which lowers fuel-injection pressure and engine performance, and increases particulate emissions. Typical European diesel fuels contain about 250,000 particulates per 100 ml. Particulate sizes of 4 μm to 7 μm are particularly critical. High-performance fuel filters with high filtration efficiency are therefore necessary to prevent damage caused by particulates.

Water in diesel fuel

Up to approximately 100 mg/kg water dissolve in diesel fuel at room temperature. The solubility limit depends on the composition of the diesel fuel, particularly the aromatic-compound content, the additives, the proportion of biodiesel, and the ambient temperature. The standard EN 590 permits a maximum water content of 200 mg/kg. Although diesel fuels in storage tanks can contain much higher amounts of water, market surveys of fuels show that water content rarely exceeds 200 mg/kg. Samples often do not detect any water, or detection is incomplete, since water is deposited on walls in a separate phase in the form of undissolved, "free" water, or it settles at the bottom. Whereas dissolved water does not damage the fuel-injection system, very small quantities of free water can cause wear or corrosion damage to fuel-injection components within a short period of time.

Additives

Additives, long a standard feature in gasolines, have attained increasing significance as quality improvers in diesel fuels. Additive packages, which are effective in a number of respects, are used for the most part. As the total concentration of the additives generally lies below 0.1 %, the fuel's physical characteristics – such as density, viscosity, and boiling curve – remain unchanged.

Lubricity enhancers

It is possible to improve the lubricity of diesel fuels with poor lubrication properties, caused, for example, by hydration processes during desulfurization, by adding fatty acids or glycerides. Biodiesel also contains glycerides as a byproduct. In this case, if diesel fuel already contains a proportion of biodiesel, no further lubricity enhancers are added.

Cetane improvers

Nitric-acid esters of alcohols, which shorten the duration until ignition (ignition lag), are frequently used as cetane improvers. They help, particularly during cold starting, to prevent an increase in combustion noise (engine noise) and extreme smoking.

Flow improvers

Flow improvers consist of polymer substances that lower the filterability limit. They are generally added in winter to ensure trouble-free operation at low temperatures. Although flow improvers cannot prevent the precipitation of paraffin crystals from diesel fuel, it can severely limit their growth. The crystals produced are then so small that they do not clog the filter pores completely and allow the passage of fuel until they melt again as the engine heat begins to build up.

Detergents

Detergents are cleaning additives which are added to the fuel to keep nozzle holes free from deposits on the inside and at the outlet. The cleaning effect of detergents is based on their molecular structure of groups of different polarity: Long nonpolar hydrocarbon chains ensure good solubility in the fuel while coordination to the for the most part also polar coatings is facilitated by way of polar groups. Detergents function in this way as carriers with which existing deposits can be removed via the fuel. Furthermore, additives can also prevent the build-up of new deposits by coordination of the polar groups to metal surfaces.

Corrosion inhibitors

Typical corrosion inhibitors are longer-chain organic monomeric and dimeric acids, amines or ammonium salts. These compounds are deposited with their polar end on the surfaces of metal parts and protect them against corrosion if water is entrained.

Antifoaming agents (defoamers)

Adding defoamers helps to avoid excessive foaming when the vehicle is refueled quickly. Antifoaming agents such as silicones reduce the surface tension, causing bubbles that are created to burst rapidly.

Blend components for fossil diesel fuels

Biodiesel

At present, biodiesel is the most important alternative fuel for diesel engines. The term biodiesel covers fatty acid esters which are created through transesterification of oils or greases with methanol. Fatty acid methyl ester (FAME) is created. The molecules of biodiesel are, in terms of size and properties, far more similar to diesel fuel than to vegetable oil. Therefore, biodiesel cannot under any circumstances be equated with vegetable oil. Nevertheless, biodiesel differs significantly from petroleum diesel, since the fatty acid esters are polar and chemically reactive. Conventional diesel fuel on the other hand is an inert and nonpolar mixture of paraffins and aromatic compounds.

Production

Vegetable oils or animal fats can be used as the starting material for biodiesel. In Europe primarily rape oil is used, in North and South America soybean oil, in Asia palm oil, and on the Indian subcontinent to an even smaller extent the oil from the purgier nut (jatropha). Used frying oil methyl ester (UFOME) is produced worldwide. Because of the global trade in biodiesel and its raw materials, fuels containing FAME as a rule contain mixtures from different sources.

Because transesterification of grease and oils can be technically carried out more easily with methanol than it can with ethanol, the methyl esters are manufactured almost exclusively. Methanol is generally produced from coal. Therefore, fatty acid methyl esters cannot, strictly speaking, be seen as fully biogenic.

Properties

The properties of biodiesel are determined by various factors. The various vegetable oils differ in the composition of the fatty-acid blocks and demonstrate typical fatty-acid patterns. The type and quantity of unsaturated fatty acids have, for example, a decisive influence on the cold resistance and resistance to aging of biodiesel. The properties are also affected by the pretreatment of the vegetable oil and the production process of the biodiesel.

Standards

The quality of biodiesel is regulated in fuel standards. If technically tenable, limitations with regard to raw materials must be avoided. The quality requirements for biodiesel are therefore described not by way of the raw-material composition, but predominantly by way of the material properties. It is essential in particular to ensure sufficient cold resistance, good resistance to aging (oxidation stability), and to eliminate contamination caused by the process.

The European standard EN 14214 is the most internationally comprehensive specification for biodiesel and describes the requirements relating to the use of biodiesel as a pure fuel and as a blend component for diesel fuel. Good-quality biodiesel is defined for both areas of application (Table 3).

The American biodiesel standard ASTM D6751 is less quality-oriented. For example, the minimum requirement for oxidation stability in this standard is less than half the aging reserve (induction period) specified in EN 14214. Thus, in the USA the risk of problems arising as a result of fuel aging, in particular under limit-value application and field conditions, is higher.

Other countries such as Brazil, India and Korea have geared themselves, to a large extent, towards the European B 100 standard EN 14214.

Use in motor vehicles

Since the tax benefit was discontinued biodiesel has barely been used as a pure fuel (B 100) in Germany. The introduction of diesel particulate filters and more stringent requirements regarding vehicle emissions and compliance over high mileages under field conditions limit the use of B 100 to a few special applications. Originally B 100 was used predominantly in commercial vehicles, because the annual high mileage ensured fast consumption, which enabled problems with insufficient oxidation stability to be avoided.

From an engine viewpoint, it is more favorable to use biodiesel as a low-percentage mixture in the blend with petroleum diesel. The high proportion of petroleum diesel ensures sufficient stability and, at the same time, the biodiesel ensures a good lubricating effect.

For practical purposes, it is important to specify not only the pure component B 100, but also the diesel/biodiesel mixtures offered on the market. Here, the trend is towards admixtures up to max. 7 % biodiesel (B 7 in Europe).

In closed fleets, higher proportions of biodiesel are also used (B 30 in France, B 20 in the USA). In the case of higher contents, the high boiling point of biodiesel can cause it to be heavily introduced into the engine oil after its has been injected into the combustion chamber via

Table 3: DIN EN 14214 (June 2014): Selected requirements of FAME

Requirements	Unit	Characteristic
Density (at 15 °C), min./max.	kg/m ³	860/900
Viscosity (at 40 °C), min./max.	mm ² /s	3.5/5.0
Sulfur content, max.	mg/kg	10
Content of alkali metals (Na + K), max.	mg/kg	5.0
Content of alkaline earth metals (Ca + Mg), max.	mg/kg	5.0
Content of total glycerin, max.	% (m/m)	0.25
Acid number, max.	mg KOH/g	0.50
Oxidation stability (induction period at 110 °C), min.	hours	8
Water, max.	mg/kg	500
Total contamination, max.	mg/kg	24
CFPP ¹ , in six seasonal classes in each case, max.		
Pure fuel	°C	+5...-20 ²
Blend component	°C	+13...-10 ³
Flash point	°C	> 101

¹ Filterability limit.

² Determined nationally, for Germany 0 to -20 °C.

³ Determined nationally, for Germany 0 to -10 °C.

condensation on the cylinder walls. This affects, above all, vehicles which are fitted with diesel particulate filters, and in which regeneration occurs by way of a retarded secondary injection. Depending on the application, it is possible, particularly in slow part-load operation, to encounter an unacceptably high introduction of biodiesel, which necessitates shorter oil-replacement intervals.

The higher boiling point of biodiesel can be taken into account by application in new developments of engines and vehicles. Consequently, a new norm has been developed in Europe for B20/B30, EN 16709 [16], in which two grades with different biodiesel content are specified: B20 with 14 to 20 % and B30 with 24 to 30 % biodiesel. To avoid problems caused by fuel aging, the minimum oxidation-stability requirement from EN 590 (B7) has been adopted for both grades, i.e. the induction time must be at least 20 hours, irrespective of the biodiesel content. In this way, the new European standard EN 16709 differs fundamentally from the American standard ASTM D 7467 [17], in which merely an induction time of at least 6 hours was laid down for biodiesel contents of 6 to 20 %.

The standard EN 16734 [18] has also been adopted in Europe; this standard covers diesel fuels with biodiesel contents of 0 to 10 % and thus overlaps to a large extent with the standard EN 590 for standard diesel fuel (0 to 7 % biodiesel). Still no market demand for this additional grade is to be discerned. Furthermore, this fuel is not permitted to be placed on the market in all European countries.

Vegetable oil

Vegetable oil which has not been transesterified into biodiesel may not be admixed with diesel fuels due to its high density, high viscosity, and non-volatility. Previously, vegetable oil, particularly rape oil, was used with great success in older diesel engines subject to minimal emission requirements. Today, vegetable oils can only be used in a few modern diesel engines with high-pressure fuel injection, for example in tractors which have been approved for vegetable-oil operation.

Paraffinic diesel fuels

Further blend components for fossil diesel fuels are purely paraffinic fuels, which consist exclusively of saturated hydrocarbons. Since there are no aromatic components, the particulate, HC and CO emissions of purely paraffinic fuels are significantly reduced.

Paraffinic fuels can be created in three different ways:

- Fischer-Tropsch process [19]
- Hydrogenation of vegetable oils
- COD process (Conversion of Olefins to Distillates)

Fischer-Tropsch process

The starting product required is synthesis gas, which consists of hydrogen and carbon monoxide. It can be created from natural gas, coal or biomass. By converting synthesis gas, it is possible to build up on catalysts linear, straight-chain hydrocarbons (n-paraffins). The Fischer-Tropsch catalysts function quite non-specifically such that a wealth of different components is obtained, starting with gases through short-chain gasoline components, kerosene and diesel paraffins, right down to oils and waxes of high molecular weight. For reasons of economy, splitting the production mixture is for the most part optimized to a maximum diesel yield. The fuels obtained in accordance with this process are known as synthetic diesels.

These fuels were originally also referred to as designer fuels, because the notion existed that the composition of synthetic diesel fuel could be geared exactly to the demands of diesel-engine technology. In view of the wide range of products obtained from Fischer-Tropsch synthesis, the notion of producing fuels of customized composition no longer appears realistic.

The terms Gas-to-Liquid (GtL), Coal-to-Liquid (CtL) and Biomass-to-Liquid (BtL) are commonly used, depending on whether the paraffins have been created from natural gas, coal or biomass.

The production of CtL and GtL is economically significant. The production of GtL is only worthwhile with large natural-gas deposits where the natural gas cannot be attributed to any direct use.

CtL and GtL are based on fossil energy sources such that no reduction in CO₂ emissions is achieved. In the case of BtL, there is no CO₂ advantage. However, industrial plants that produce BtL are not yet in operation.

In this approach of using biomass completely, plants are first broken down into chemical building blocks from which fuels are then synthesized in a subsequent step. Fuels which are manufactured in this way are called second-generation fuels. This procedure differs fundamentally from the hitherto customary methods, which transform existing components in agricultural crops such as fats, starches and sugars through chemical or enzymatic separation (transesterification, fermentation) into fuels (biodiesel or bioethanol) (first-generation fuels).

Hydrogenation of vegetable oils
Paraffinic diesel fuels can also be obtained by hydrogenating fats and oils. Unlike transesterification into biodiesel, conversion with hydrogen places fewer demands on the origin and quality of the starting materials. Hydrogenation results in a cracking of fats and oils, during which all the oxygen atoms and unsaturated bonds are also removed. Long-chain paraffins are obtained from the fatty acids, the glycerin content is converted into propane gas, and the oxygen is bound as water. Because paraffins are created from

biomass in this way, they are referred to as bioparaffins. The production of bioparaffins can be implemented in separate plants or even integrated into existing refinery processes.

The hydrogenation of vegetable oils is developing increasingly into a significant alternative to the production of biodiesel.

COD process
The third way of producing paraffins is to convert olefins in accordance with the COD process (Conversion of Olefins to Distillates), frequently only as a subsequent step to previous refinery processes. Olefinic product fractions are used for this purpose. Paraffins can be produced by oligomerization, a specifically controlled reaction between individual olefins, and hydrogenation.

Properties of paraffinic diesel fuels
Regardless of the type of production, paraffinic hydrocarbon mixtures with very similar chemical compositions and excellent engine properties are created. The fuels are sulfur- and aromatic-free and have partly high cetane numbers. But because their density is below the lower limit value defined in EN 590, the new specification EN 15940 [20] was developed (Table 4) which describes a quality of the fuel for pure use and limits use to such vehicles which are explicitly intended for this fuel. The three manufacturing processes de-

**Table 4: DIN EN 15940 (September 2016):
Selected requirements of paraffinic diesel fuels**

Requirements		Unit	Characteristic
Cetane number, min.	Class A	–	70
	Class B	–	51
Density (at 15 °C), min./max.	Class A	kg/m ³	765/800
	Class B	kg/m ³	780/810
Viscosity (at 40 °C), min./max.		mm ² /s	2.0/4.5
Sulfur content, max.		mg/kg	5
Total aromatic content (including polyaromatics), max.		% (m/m)	1.1
Lubricity, “wear scar diameter”, max.		µm	460
FAME content, max.		% (V/V)	7
Water, max.		mg/kg	200
Total contamination, max.		mg/kg	24
CFPP ¹ in six seasonal classes, max. ²		°C	+5 to –20
Flash point		°C	>55

¹ Filterability limit.
² Determined nationally, for Germany 0 to –20 °C.

scribed require an additional isomerization step in which linear molecules are converted into branched molecules. The quality requirements defined in EN 15940, particularly the required cold resistance, are achieved in this way.

Use in motor vehicles

Pure paraffinic fuels are used mostly in older vehicles with no diesel particulate filters and lend themselves to use above all in centers of population in order to reduce the particulate load on a local level. Paraffinic hydrocarbons are ideal for marketing as blend components in premium diesel fuels. Furthermore, diesel fuels which fail to satisfy the requirements established in EN 590 can be improved by the addition of paraffinic components to such an extent that they conform to the standard.

Natural gas

Fossil natural gas

The main constituent of natural gas is methane (CH_4) with a minimum content of 80 % (V/V). Further constituents are inert gases, such as carbon dioxide, nitrogen and short-chain hydrocarbons. Oxygen and hydrogen are also contained. Natural gas is promoted worldwide and can be prepared with relatively little effort. Depending on its origin, however, its composition varies, which results in fluctuations when it comes to density, calorific value and knock resistance. How natural gas is transported also influences its composition.

Unlike when it is supplied in gas form through pipelines, natural gas that is distributed by ships in liquid form as LNG (liquefied natural gas) contains hardly any CO_2 , pentane, hexane, and hydrogen sulfide. The properties of natural gas as a fuel are defined for Germany in the standard DIN 51624 [21]. There is also a European standard, EN 16723-2 [22], in which standardized quality requirements with regard to natural gas, to pure biomethane and to blends of natural gas and biomethane are defined.

Biomethane

Biomethane is treated biogas which is made from biogenic materials such as energy crops, liquid manure or waste containing biomass. The CO_2 savings from the use of biogenic basic material can be set off against the CO_2 emissions during combustion in such a way that the total CO_2 emissions for biomethane are clearly reduced compared with fossil natural gas.

Storage of natural gas

Natural gas is stored either in gas form (CNG, compressed natural gas) at a pressure of 200 bar or as a liquefied gas (LNG, liquefied natural gas) at -162°C in a cold-resistant tank. LNG requires only one third of the storage volume of CNG, however the storage of LNG requires a high expenditure of energy for liquefying. Filling stations dispense natural gas almost exclusively in the form of CNG. LNG is as a rule only offered if the filling station can already be supplied with LNG.

CO₂ emissions

The hydrogen/carbon ratio of natural gas is roughly 4:1, that of gasoline on the other hand is 2.3:1. As a result of the lower amount of carbon in natural gas, it produces less CO₂ and more H₂O than gasoline when it is burned. A spark-ignition engine converted to natural gas creates, for a comparable power output without any further optimization, approximately 25 % fewer CO₂ emissions than when run on gasoline. Since direct methane emissions have 20 times greater climate-damaging potential than CO₂, it is necessary to take into account in the overall footprint a small methane slip in the event of incomplete combustion. Even dispensing gaseous methane from LNG tanks is relevant to the carbon footprint when an overpressure builds up and this is discharged to the atmosphere.

Use of natural gas

Natural gas is used in vehicles with spark-ignition engines. Passenger cars for operation on natural gas are for the most part bivalent and also permit, separately from the use of natural gas, the use of gasoline. Buses with spark-ignition gas engines run in monovalent mode on natural gas. In compression-ignition gas engines natural gas is burned in mixed mode with diesel fuel. Whether the market for natural-gas vehicles will gain in importance depends primarily on the purchase and running costs for the vehicle owner, but also on the expansion of the filling-station infrastructure to CNG and LNG.

Liquefied petroleum gas

Liquefied petroleum gas (LPG), also called liquid gas, is obtained during the extraction of crude oil and during different refinery processes. LPG is a mixture of the main components of propane and butane. It can be liquefied at room temperature under comparatively low pressure. Because it has a lower carbon content than gasoline, approximately 10 % less CO₂ is produced during combustion. The octane number is approximately 100 - 110 RON. The demands placed on LPG for use in motor vehicles are laid down in the European standard EN 589 [23].

Passenger cars for liquefied petroleum gas are also for the most part bivalent and can therefore be run separately on gasoline and liquefied petroleum gas. However, liquefied petroleum gas can cause significant wear on the intake valves if valve or seat-insert materials have not been adapted to running on LPG.

Hydrogen

Production

Hydrogen can be created by chemical processes from natural gas, coal, crude oil or biomass, and by electrolysis from water. Today hydrogen is predominantly obtained on an industrial scale by steam reformation from natural gas. In this process, CO_2 is released in such a way that the use of hydrogen does not necessarily result in a CO_2 advantage compared with gasoline, diesel or the direct use of natural gas in the internal-combustion engine.

A reduction in CO_2 emissions is achieved then when hydrogen is regeneratively created from biomass or by electrolysis from water provided that regeneratively generated current is used for this purpose. No CO_2 emissions occur locally when hydrogen is burned.

Storage

Hydrogen may have a very high mass-related energy density (approximately 120 MJ/kg and thus almost three times as high as that of gasoline), but its volume-related energy density is very low on account of the low specific density. When it comes to storage, this means that the hydrogen has to be compressed either under pressure (at 350 – 700 bar) or by liquefaction (cryogenic storage at -253°C) in order to achieve an acceptable tank volume. Another possibility is for the hydrogen to be stored as a hybrid. Certain metals and alloys form, with hydrogen, metal hydrides. The hydrogen that is embedded in the metal's atomic lattice can be released again by heating.

Use in motor vehicles

Hydrogen can be used both in fuel-cell drives and directly in internal-combustion engines. In the long term, emphasis is expected to be placed on use in fuel cells because this achieves better efficiency than in hydrogen internal-combustion engines.

e-fuels

The term e-fuels is used for electrofuels and comprises fuels which are constituted using electric current. In the first step, water is broken down into hydrogen and oxygen in an electrolysis process. The created hydrogen, e-H_2 , is the simplest e-fuel, which can be used directly, for example in a fuel cell.

Alternatively, methane, e-CH_4 , can also be produced from e-H_2 , where e-H_2 is converted with CO_2 . Liquid e-fuels such as e-diesel consist of paraffins which can be constituted from e-H_2 and carbon monoxide by means of Fischer-Tropsch synthesis. Carbon monoxide is available from CO_2 by a reversal of the water-gas shift reaction. All these e-fuels (e-H_2 , e-CH_4 , e-diesel) do not differ from the analogous fossil or biogenic components or are qualitatively very similar to these.

Although the production of e-fuels requires an additional energy expenditure which is above its own energy content and the costs of their production, depending on the expansion stage, far exceed the costs of the fossil components, e-fuels will be given opportunities in the future. Without e-fuels, for which a significant reduction of the production costs is forecast, the long-term CO_2 targets in the traffic sector cannot be achieved.

Ethers

Dimethyl ethers

Dimethyl ether (DME) with the structure $\text{H}_3\text{C}-\text{O}-\text{CH}_3$ has a boiling point of -25°C and must therefore be kept under pressure at room temperature so that it can be injected in liquid form. The methyl groups in the DME are isolated by oxygen bridges. Consequently, little soot is produced during combustion. DME also has a high cetane number, reducing combustion and lowering NO_x emissions. The high oxygen content of DME however gives rise to a low calorific value H_u of 28.8 MJ/kg and thus high volumetric increased consumption compared with diesel fuel (H_u of 43.0 MJ/kg).

Although DME from a combustion perspective is an ideal fuel for diesel engines, in reality only a few specially adapted vehicles use this fuel. DME liquefied under pressure is, due to its high vapor pressure and its low viscosity, not compatible with the existing injection components and engine designs developed for conventional diesel fuels. The use of DME will, due to the necessary additional technical outlay on the vehicle, remain limited to just a few fleet vehicles in niche markets, even though the synthesis gas required to produce DME (hydrogen and carbon monoxide) can be manufactured relatively easily from coal, natural gas, biomass or even waste materials.

Oxymethylene ethers

Oxymethylene ethers (OME) also have oxygen bridges and isolated carbon atoms. From the ethers of the structure $\text{H}_3\text{C}-\text{O}-[\text{CH}_2-\text{O}]_n-\text{CH}_3$ with $n = 1$ to $n = 5$ only the basic component with $n = 1$ from synthesis gas is up to now technically easily accessible. This oxymethylene ether, which is also called OME1 or chemically more accurately dimethoxymethane, is however not suitable either as a pure substance or as a blend component for existing diesel engines, although OME1 unlike DME is still liquid at room temperature.

With a low boiling point of 42°C , low viscosity of $0.33\text{ mm}^2/\text{s}$ and a flash point of -18°C , OME1 satisfies neither the hydraulic requirements with regard to high-pressure fuel injection nor the existing statutory safety provisions. The combustion of OME1 generates little soot. The NO_x emissions resulting from the low cetane number of 38 can be reduced via the trade-off of soot to NO_x .

The low calorific value H_u of OME1 of 22.7 MJ/kg results in a roughly 70 % higher volumetric consumption compared with diesel fuel. The higher oxymethylene ethers OME3 to OME5 have – apart from the likewise low calorific value – clearly better substance properties. Up to now, however, there are no appreciable quantities of OME3 to OME5 available extending beyond the demand for industrial use.

Table 5: Properties of liquid fuels and hydrocarbons

Material/medium	Density kg/l	Main constituents % (m/m)	Boiling temperature °C	Specific heat of evaporation kJ/kg	Net calorific value MJ/kg	Ignition temperature °C	Air requirement, theoretical kg/kg	Ignition limit Lower Upper % (V/V) gas in air
Gasoline, Regular Premium	0.720...0.775	86 C, 14 H	25...210	380...500	41.2...41.9	≈ 300	14.8	≈ 0.6
Aviation fuel	0.720...0.775	86 C, 14 H	25...210	—	40.1...41.6	≈ 400	14.7	—
Kerosene	0.77...0.83	85 C, 15 H	40...180	—	43.5	≈ 500	—	≈ 0.7
Diesel fuel	0.820...0.845	87 C, 13 H	170...260	—	43	≈ 250	14.5	≈ 0.6
		86 C, 14 H	180...360	≈ 250	42.9...43.1	≈ 250	14.5	≈ 0.6
Mineral oil (crude oil)	0.70...1.0	80...83 C, 10...14 H	25...360	222...352	39.8...46.1	≈ 220	—	≈ 0.6
Lignite tar oil	0.850...0.90	84 C, 11 H	200...360	—	40.2...41.9	—	13.5	—
Coal tar oil	1.0...1.10	89 C, 7 H	170...330	—	36.4...38.5	—	—	—
Pentane C_5H_{12}	0.63	83 C, 17 H	36	352	45.4	285	15.4	1.4
Hexane C_6H_{14}	0.66	84 C, 16 H	69	331	44.7	240	15.2	1.2
n-heptane C_7H_{16}	0.68	84 C, 16 H	98	310	44.4	220	15.2	1.1
iso-octane C_8H_{18}	0.69	84 C, 16 H	99	297	44.6	410	15.2	1
Benzene C_6H_6	0.88	92 C, 8 H	80	394	40.2	550	13.3	1.2
Toluene C_7H_8	0.87	91 C, 9 H	110	364	40.6	530	13.4	1.2
Xylene C_8H_{10}	0.88	91 C, 9 H	144	339	40.6	460	13.7	1
Ether ($C_2H_5)_2O$	0.72	64 C, 14 H, 22 O	35	377	34.3	170	7.7	1.7
Acetone ($CH_3)_2CO$	0.79	62 C, 10 H, 28 O	56	523	28.5	540	9.4	2.5
Ethanol C_2H_5OH	0.79	52 C, 13 H, 35 O	78	904	26.8	420	9	3.5
Methanol CH_3OH	0.79	38 C, 12 H, 50 O	65	1,110	19.7	450	6.4	5.5
Dimethoxymethane (OME1)	0.86	47 C, 11 H, 42 O	42	376	22.7	235	7.2	2.2
Rape oil	0.92	78 C, 12 H, 10 O	—	—	38	≈ 300	12.4	—
Rapeseed methyl ester (biodiesel)	0.88	77 C, 12 H, 11 O	320...360	—	36.5	283	12.8	—

Viscosity at 20 °C in mm²/s (= cSt): gasoline ≈ 0.6; ethanol ≈ 1.5; methanol ≈ 0.75.

Table 6: Properties of gaseous fuels and hydrocarbons

Material/medium	Density at 0 °C and 1,013 mbar kg/m ³	Main constituents % (m/m)	Boiling temperature at 1,013 mbar °C	Net calorific value		Ignition temperature °C	Air requirement, theoretical kg/kg	Ignition limit	
				Fuel MJ/kg	Air-fuel mixture MJ/m ³			Lower	Upper
Liquid petroleum gas	2.25 ⁽¹⁾	C ₃ H ₈ , C ₄ H ₁₀	-30	46.1	3.39	≈ 400	15.5	1.5	15
Municipal gas	0.56...0.61	50 H ₂ , 8 CO, 30 CH ₄	-210	≈ 30	≈ 3.25	≈ 560	10	4	40
Natural gas H (North Sea)	0.83	87 CH ₄ , 8 C ₂ H ₆ , 2 C ₃ H ₈ , 2 CO ₂ , 1 N ₂	-162 (CH ₄)	46.7	—	584	16.1	4.0	15.8
Natural gas H (Russia)	0.73	98 CH ₄ , 1 C ₂ H ₆ , 1 N ₂	-162 (CH ₄)	49.1	3.4	619	16.9	4.3	16.2
Natural gas L	0.83	83 CH ₄ , 4 C ₂ H ₆ , 1 C ₃ H ₈ , 2 CO ₂ , 10 N ₂	-162 (CH ₄)	40.3	3.3	≈ 600	14.0	4.6	16.0
Water gas	0.71	50 H ₂ , 38 CO	—	15.1	3.10	≈ 600	4.3	6	72
Blast-furnace gas	1.28	28 CO, 59 N ₂ , 12 CO ₂	170	3.20	1.88	≈ 600	0.75	≈ 30	≈ 75
Sewage gas (manure gas) ²⁾	—	46 CH ₄ , 54 CO ₂	—	27.2 ⁽³⁾	3.22	—	—	—	—
Hydrogen H ₂	0.090	100 H ₂	-253	120.0	2.97	560	34	4	77
Carbon oxide CO	1.25	100 CO	-191	10.05	3.48	605	2.5	12.5	75
Methane CH ₄	0.72	75 C ₂ , 25 H ₂	-162	50.0	3.22	650	17.2	5	15
Acetylene C ₂ H ₂	1.17	93 C ₂ , 7 H ₂	-81	48.1	4.38	305	13.25	1.5	80
Ethane C ₂ H ₆	1.36	80 C ₂ , 20 H ₂	-88	47.5	—	515	17.3	3	14
Ethene C ₂ H ₄	1.26	86 C ₂ , 14 H ₂	-102	14.1	—	425	14.7	2.75	34
Propane C ₃ H ₈	2.0 ⁽¹⁾	82 C ₃ , 18 H ₂	-43	46.3	3.35	470	15.6	1.9	9.5
Propene C ₃ H ₆	1.92	86 C ₃ , 14 H ₂	-47	45.8	—	450	14.7	2	11
Butane C ₄ H ₁₀	2.7 ⁽¹⁾	83 C ₄ , 17 H ₂	-10; +1 ⁽³⁾	45.6	3.39	365	15.4	1.5	8.5
Butene C ₄ H ₈	2.5	86 C ₄ , 14 H ₂	-5; +1 ⁽³⁾	45.2	—	—	14.8	1.7	9
Dimethyl ether C ₂ H ₆ O	2.05 ⁽⁴⁾	52 C ₂ , 13 H ₂ , 35 O	-25	28.8	3.43	235	9.0	3.4	18.6

¹ Density of liquified gas 0.54 kg/l, density of liquid propane 0.51 kg/l, density of liquid butane 0.58 kg/l.² Purified sewage gas contains 95 % CH₄ (methane) and has a calorific value of 37.7 MJ/kg.³ First value for isobutane, second value for n-butane and n-butene.⁴ Density of liquified dimethyl ether 0.667 kg/l.

References

- [1] DIN EN 228:2017, Automotive fuels – Unleaded petrol – Requirements and test methods.
- [2] ASTM D4814-16b, Standard Specification for Automotive Spark-Ignition Engine Fuel.
- [3] Standard draft E DIN 51626-1:2010-04, Automotive fuels – Requirements and test methods – Part 1: Petrol E 10 and petrol E 5.
- [4] DIN EN ISO 5164:2014, Petroleum products – Determination of knock characteristics of motor fuels – Research method (ISO 5164:2014).
- [5] DIN EN ISO 5163:2014, Petroleum products – Determination of knock characteristics of motor and aviation fuels – Motor method (ISO 5163:2014).
- [6] DIN EN 13016-1:2007, Liquid petroleum products – Vapor pressure – Part 1: Determination of air saturated vapor pressure (ASVP) and calculated dry vapor pressure equivalent (DVPE).
- [7] DIN CEN/TS 15293:2011, Automotive fuels – Ethanol (E 85) automotive fuel – Requirements and test methods.
- [8] ASTM D5798-13a, Standard Specification for Ethanol Fuel Blends for Flexible-Fuel Automotive Spark-Ignition Engines.
- [9] DIN EN 590:2017, Automotive fuels – Diesel fuel – Requirements and test methods.
- [10] DIN EN 14214:2014, Liquid petroleum products – Fatty acid methyl esters (FAME) for use in diesel engines and heating applications – Requirements and test methods.
- [11] ASTM D975-15c, Standard Specification for Diesel Fuel Oils.
- [12] ASTM D6751-15c, Standard Specification for Biodiesel Fuel Blend Stock (B100) for Middle Distillate Fuels.
- [13] DIN EN ISO 5165 (prEN ISO 5165:2016), Petroleum products – Determination of the ignition quality of diesel fuels – Cetane engine method (ISO 5165:1998).
- [14] DIN EN ISO 12156-1:2016, Diesel fuel – Assessment of lubricity using the high-frequency reciprocating rig (HFRR) – Part 1: Test method.
- [15] DIN EN 15751:2014, Automotive fuels – Fatty acid methyl ester (FAME) fuel and blends with diesel fuel – Determination of oxidation stability by accelerated oxidation method.
- [16] DIN EN 16709:2015, Automotive fuels – High FAME diesel fuel (B20 and B30) – Requirements and test methods.
- [17] ASTM D7467-15c, Standard Specification for Diesel Fuel Oil, Biodiesel Blend (B6 to B20).
- [18] DIN EN 16734:2016, Automotive fuels – Automotive B10 diesel fuel – Requirements and test methods.
- [19] L. König, J. Gaube: Fischer-Tropsch-Synthese, Neuere Untersuchung und Entwicklungen. Article in Chemie Ingenieur Technik, Volume 55/1, 1983.
- [20] DIN CEN/TS 15940:2016, Automotive fuels – Paraffinic diesel fuel from synthesis or hydrotreatment – Requirements and test methods.
- [21] DIN 51624: 2008, Automotive fuels – Compressed natural gas – Requirements and test methods.
- [22] DIN EN 16723-2:2017, Natural gas and biomethane for use in transport and biomethane for injection in the natural gas network – Part 2: Automotive fuels specification.
- [23] DIN EN 589:2012, Automotive fuels – LPG – Requirements and test methods.

Urea/water solution

Application

A urea/water solution is used to reduce nitrogen oxides (NO_x) in the catalytic exhaust-gas treatment of diesel engines with the aid of SCR systems (see SCR system). To this end, this solution is not used as a fuel additive, but metered directly into the exhaust-system branch.

Designations

The solution used under the protected trade name AdBlue® consists of 32.5 % technically pure urea and 67.5 % demineralized water. The chemical total formula is: $\text{CH}_4\text{N}_2\text{O}$. Further designations are:

- carbamide in aqueous solution,
- Aqueous Urea Solution 32.5 % (AUS 32),
- NO_x reduction agent in acc. with DIN 70070 [1] or ISO 22241 [2] Part 1 to Part 4,
- urea/water solution,
- Diesel Exhaust Fluid (DEF),
- Agente Reductor Líquido de Óxido de Nitrogénio Automotivo (ARLA 32).

Properties

High temperatures and direct sunlight are to be avoided due to decomposition by heat. When AdBlue is heated above 25 °C, an ammonia solution is created as an aqueous solution or as an aerosol of ammonia as a hazard-determining substance.

AdBlue attacks base metals. Copper, alloys containing copper and unalloyed and galvanized steels, aluminum and glass are not suitable for contact with a urea/water solution.

AdBlue has a strong tendency to crystallization. White crystals form when the water that it contains evaporates. The growth of crystals is encouraged by increased air moisture. AdBlue has high penetration ability and a strong dispersion capacity. Lines/pipes can be blocked by crystallization. Blockages can be dissolved by warm water. It is essential to look out for the growth of crystals along electrical wiring/leads. Electric contacts can be jumpered by the high conductivity. Leads containing copper are corroded.

The product is not classified as hazardous and does not have to be specially labeled. The substance is not irritant, but after repeated or prolonged skin contact can cause degreasing of the skin and dermatitis. Experience in practice has shown reddened skin after a few minutes in contact with the medium.

With water hazard class 1, AdBlue is classified as slightly hazardous to water and may not be released into the environment.

At ambient temperatures below -10 °C containers, lines and equipment must be fitted with cold insulation and heating. Frozen AdBlue incurs no damage and can therefore be reused once it has thawed.

Storing AdBlue at temperatures below 0 °C is to be avoided due to the risk of freezing and the associated volumetric expansion. There is the danger of the container splitting. A leak may only become visible after thawing.

References

- [1] DIN 70070: Diesel engines – NO_x reduction agent AUS 32 – Quality requirements.
 [2] ISO 22241: Diesel engines – NO_x reduction agent AUS 32 – Quality requirements.

Table 1: Properties of AdBlue

Delivery form	Aqueous solution
Appearance	Clear and colorless
Odor	Odorless up to slight odor of ammonia
Solubility	Mixable with water in any ratio, poorly soluble in aliphatic hydrocarbons
Congeeing point	-11.5 °C
Reaction	Weakly alkaline
Electric conductivity	Liquid and crystals good electric conductivity

Refrigerants for air-conditioning systems

The core component of the air-conditioning system in motor vehicles is as a rule a cold vapor-compression refrigerating system (see Passenger-compartment climate control). The underlying thermodynamic cyclic process – the Carnot cycle – is approximated with this system (see Thermodynamics). The two isentropic changes of state (according to Carnot) are approximated by a air compressor and a throttle element and the isotherms are approximated by evaporating and condensing refrigerant. Operating fluids for which the phase changes occur at tenable pressures are required especially for evaporation and condensation.

Table 1 shows the property values of the refrigerants R134a and R1234yf. The global warming potential (GWP) is a measure of the relative contribution of a greenhouse gas to the greenhouse effect. The reference forms the CO₂ with the global warming potential of one. The emission of 1 kg R134a would accordingly have a greenhouse effect comparable with the emission of 1300 kg CO₂.

Table 1: Property values of the refrigerants R134a and R1234yf

Property	R134a	R1234yf
Boiling point	-26 °C	-29 °C
Vapor pressure at 25 °C	6.56 bar	6.64 bar
Vapor pressure at 80 °C	25.97 bar	24.38 bar
Vapor density	32.4 kg/m ³	37.6 kg/m ³
GWP	approx. 1300	approx. 4

Development up to today

Refrigerant R12

Initially, combustible substances (e.g., methyl chloride CH₃Cl, sulfur dioxide SO₂ or diethyl ether C₄H₁₀O) were used in refrigerant systems. In around 1930 the American engineer Thomas Midgley developed chlorofluorocarbons (CFCs). The most important substance from this family for motor vehicles was R12 (CCl₂F₂). This substance accepted as non-poisonous and non-combustible was used in practically all passenger-car air-conditioning systems and was additionally available for a great many other applications.

In 1974 the chemists Molina and Rowland put forward the hypothesis that substances containing chlorine damage the ozone layer that protects us [1]. Subsequent observations of the ozone hole over Antarctica bore out this hypothesis. The evaluation of precise measurements finally provided certainty. This resulted on 16 September 1987 (in Montreal) in the signing of the Montreal Protocol, in which the resolution was passed to renounce chlorine chemistry.

Refrigerant R134a

From around 2000 onwards, the refrigerant R134a (C₂H₂F₄) was used as a substitute substance. But heightened sensitivity to all environmental impacts has also led to the limitation or restriction on the use of R134a. According to an EU Directive [2], air-conditioning systems designed to contain fluorinated greenhouse gases with a global warming potential of over 150 are to be prohibited in motor vehicles. R134a has a global warming potential of around 1300 and therefore falls below this Directive. Since January 2011 new vehicle types have no longer been permitted to be equipped with R134a and since January 2017 all new cars have been required to be equipped with a refrigerant that has a global warming potential of less than 150. Also applicable aside from this Directive is the EC Regulation [3], the main purpose of which is the avoidance or reduction of the emission of fluorinated greenhouse gases.

Refrigerant R1234yf

At the moment the only new refrigerant available that is produced in sufficient quantities is R1234yf ($\text{C}_3\text{H}_2\text{F}_4$) with a global warming potential of 4. Since 2017 it has, apart from a very few exceptions, been used by all manufacturers in all passenger-car air-conditioning systems. The very low global warming potential chiefly results from the fact that this substance has a very short service life in the atmosphere. R1234yf is however not undisputed. It is more easily flammable than R134a and in the event of a fire caustic and toxic reaction products are created. Therefore automobile manufacturers take additional safety precautions to ensure the customarily high safety of occupants.

Refrigerant CO_2

Air-conditioning systems with CO_2 as refrigerant are used in some upper-category passenger cars. CO_2 has by definition a global warming potential of 1 and is thus naturally predestined to be used. Unfortunately a CO_2 air-conditioning system cannot be operated with the same process and the same components. The process works on the high-pressure side in the supercritical range at very high pressures (100 to 150 bar). Therefore all the components must be newly developed and dimensioned for these unusually high pressures. Since such new developments affect not only development, production and sales but also to a significant extent service, a comprehensive market launch on 1st January 2017 could not be achieved.

An additional risk also arises in CO_2 systems. Because of the extremely high pressures a considerable amount of energy (pressure times volume) is stored in the system. This energy could for example in an accident be additionally released and cause further damage. Therefore enhanced safety precautions must also be taken for CO_2 air-conditioning systems.

Alternative possibilities

It is safe to assume that R1234yf will not be the last refrigerant used. One possibility is to build on the experiences with the CO_2 air-conditioning systems and expand this technology as a long-term solution. If this does not prove successful, natural substances which are really neutral in their impact on the environment can be considered. There have already been studies in this connection at the time of the discussions regarding renunciation of chlorine chemistry when scientists were looking for substitute substances for refrigerants containing chlorine [4]. The most natural substances would be water and air.

Water as refrigerant

In principle, water could be used in a cold vapor-compression refrigerating system, but evaporation and condensation would take place at extremely low pressures. In other words, the tiniest leaks would raise the pressure and the system function would no longer be guaranteed. The function has been verified in the laboratory. However, this would not be easily feasible at the current time.

Air as refrigerant

Air could be used in systems which operate according to the Joule process. Such systems are installed in virtually all commercial aircraft. However, the design of systems with much smaller output figures – i.e. for passenger-car applications – is more difficult and at the least entails diminished effectiveness. The process has been used with a certain degree of success for climate control in rail vehicles, for example in the ICE 1. The system output here is on a comparable scale to aircraft, but conventional R134a systems have a significantly better effectiveness.

At the moment the German Federal Environment Agency is supporting a research project in which systems with natural refrigerants are to be reassessed [5]. Since the Joule process is a pure gas process, it is inferior to the Carnot process in terms of effectiveness. It is used economically in aircraft due to the fact that some positive supplementary effects

are utilized. The air-conditioning system here has the additional task of pressurizing the cabin. The system weighs less than half as much as a comparable cold vapor-compression system and the drive can be effectively realized with compressed air from the engines.

Combustible substances as alternative refrigerants

Combustible substances could be further alternatives. In refrigerators R12 for example has been superseded by isobutane. In motor vehicles, however, the situation is a little more difficult in that as the result of a leak for example in the evaporator combustible substances would inevitably get into the passenger cell.

References

- [1]: M.J. Molina; F.S. Rowland: Stratospheric sink for chlorofluoromethanes: chlorine atom catalysed destruction of ozone. *Nature*, No. 249 pp 810–812, June 28, 1974.
- [2] Directive 2006/40/EC of the European Parliament and of the Council of 17 May 2006 relating to emissions from air-conditioning systems in motor vehicles and amending Council Directive 70/156/EEC.
- [3] Regulation (EC) No. 842/2006 of the European Parliament and of the Council of 17 May 2006 on certain fluorinated greenhouse gases.
- [4] H. Kruse: Alternative Kälteprozesse unter Umweltschutzgesichtspunkten – DKV status report of the Deutscher Kälte- und Klimatechnischer Verein No. 3 (1989).
- [5] Umweltbundesamt: Erprobung, Messung und Bewertung von Systemen mit natürlichen Kältemitteln zum nachhaltigen Kühlen und Heizen von öffentlichen Verkehrsmitteln – Ersatz fluorierter Treibhausgase. Duration 07 May 2015 – 31 March 2018.

Tolerances

Form deviations

No component can be manufactured with absolute precision, i.e. with an ideal geometry. A whole range of causes is responsible for this, e.g., deformations as a result of machining forces, the play in guides of machine tools, tool wear, vibrations, and much more. The consequence of this is that every component exhibits dimensional, geometrical and surface deviations. To guarantee the function it is necessary that all the permissible deviations be indicated in the drawing.

Dimensional deviations

Each dimension can only be manufactured between an upper and a lower limit dimension. The difference between these two limits is called the dimensional tolerance.

Geometrical deviations

Geometries also exhibit deviations from the ideal geometry. A distinction is made between form and positional deviations.

Surface deviations

The surface of a workpiece likewise cannot be manufactured to the ideal smoothness levels. The description of the surface texture can be specified with the aid of the primary, waviness and roughness profiles.

Geometrical Product Specifications

Geometrical Product Specifications (GPS) refers to all the indications in a CAD model or a drawing which are required to manufacture and check (verify) a product. Product specifications must be carried out in such a way that the component to be manufactured performs its function without limitations.

The aim of Geometrical Product Specifications is to provide a clear and complete description of the product in accordance with standardized rules. These rules are laid down in the form of GPS standards. In addition, the design intent, i.e. what is important for the function, must be apparent from a drawing.

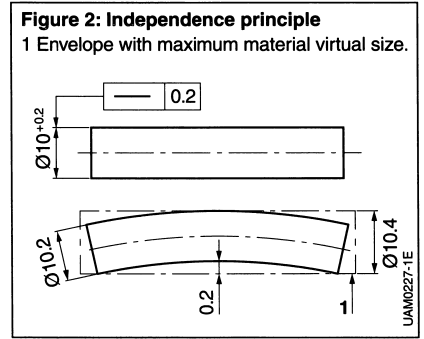
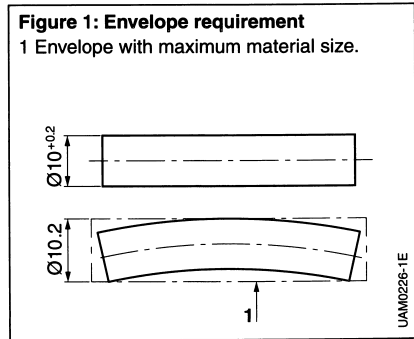
Important standards which are important for the correct interpretation of tolerances are set out in ISO 8015 [1]. These include among others the standard tolerancing principle.

Tolerancing principles

The relationship between the dimensional and geometrical tolerances is defined by the tolerancing principle. According to GPS there are two possibilities for this.

1. The envelope principle

Until 1985 no tolerancing principle was indicated on drawings, but was instead interpreted according to the Taylor principle. This corresponds to the envelope requirement at which the maximum



material size forms an ideal envelope (Figure 1). All dimensional, form and parallelism deviations must be within this envelope. This requirements can be applied to cylinders and parallel, opposing surfaces. The envelope requirement was described in 1987 in DIN 7167 [2] and established this principle as standard if nothing different was indicated in the drawing.

2. The independence principle

In 1985 ISO 8015 defined the independence principle for the first time. According to this, dimensional and geometrical tolerances are independent of each other. Adding the two tolerances produces an envelope with a maximum material virtual size (Figure 2). The entry "Tolerancing acc. to ISO 8015" is absolutely necessary for drawings to which the independence principle should apply. This principle has been the GPS standard since 2012 (after DIN 7167 was withdrawn) and is also valid without a corresponding entry in the drawing. If the envelope requirement is to apply to an individual dimension element, an $\textcircled{\text{E}}$ must be entered after the dimension. The independence principle can be canceled with the drawing entry "Dimensions acc. to ISO 14405 $\textcircled{\text{E}}$ ". In this case the envelope requirement applies to the entire drawing.

Dimensional tolerances

Dimension definition

Although dimensions are as old as the technical era, dimension characteristics were defined for the first time in 2010 in ISO 14405 [3]. According to this, different dimension characteristics can be assigned to a dimension element, e.g., a cylinder, by indicating the corresponding modification symbol after the dimension. Four different dimension characteristics and the associated modification symbols are indicated in Figure 3. The two-point dimension (local point) does not have to be indicated since without drawing indication all the dimensions have to be measured accordingly (GPS standard).

ISO tolerance system

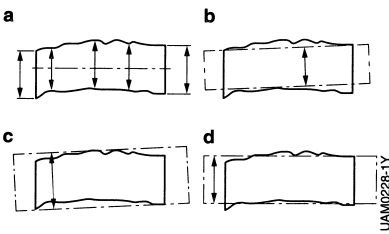
A special tolerance system has been defined for the pairing of two dimension elements in ISO 286 [4]. It applies to cylinders (cylindrical fits) and two parallel, opposing surfaces (flat fits). A special example of this is given in the chapter "Roller bearings" for bearing seats (see Roller bearings).

The tolerance class is identified by a combination of letters and code numbers. Capital letters are used for bores, lower-case letters for shafts. In this way the fundamental deviation, which indicates the position of the tolerance interval relative to the zero line, is defined (Figure 4). The code number (01 through 18) defines the fundamental tolerance degree (size of the tolerance). Both the fundamental deviation and the tolerance degree are indicated in ISO 286 in tables.

All combinations from A through z and from a through Z and all code numbers from 01 through 18 are possible as fits (e.g., $30^{+7}/30_{-6}$). Clearance, transition and interference fits can be defined in this way. To limit the multitude of possible fits, only bores with the fundamental deviation zero (code number H) are paired with different shaft tolerances. This produces the system of fits "Unit bore" (Figure 5). This is used whenever expensive tools (e.g., reamer) are used to produce bores. ISO 286 also defines a system of fits "Unit shaft". Only shafts with the fundamental deviation zero (code number h) are used. Clearance, transition and inter-

Figure 3: Dimensions acc. to ISO 14405

- a) Two-point dimension (local point) LP,
- b) Maximum inscribed dimension GX,
- c) Minimum circumscribed dimension GN,
- d) Averaged dimension (acc. to Gauss) GG.



ference fits can be achieved with these combinations.

Geometrical tolerances

Since a two-point dimension cannot define geometry properties, it additionally requires an indication with regard to the permissible form or positional deviations.

A form tolerance only defines how far the real form may deviate from the ideal form and therefore requires no reference. A surface must for example be flat within a specified tolerance, regardless of the position of this surface.

Positional tolerances limit the permissible deviations of an element from the ideal position relative to one or more elements, the so-called reference elements. A reference is always required to define a position in space. Since references are derived from real elements, they should be as precise as possible and therefore always require a form tolerance.

Six form and eleven positional tolerances (Table 1) are defined in ISO 1101 [5]. The specification for parallelism is shown in Figure 6.

General tolerances

General tolerances are used to make drawings more simple. Standard workshop precision is sufficient for all dimensions and geometry elements which are not directly important to the function or exchangeability. It can be maintained without any special effort and applies to all the dimensions and geometries which are not individually tolerated. The general tolerances are dependent on the different manufacturing processes, e.g., machining (ISO 2768, [6], [7]), metal castings (ISO 8062-3, [8]) or plastic castings (DIN 16742, [9]).

ISO 2768 is the most frequently used general tolerance and therefore indicated

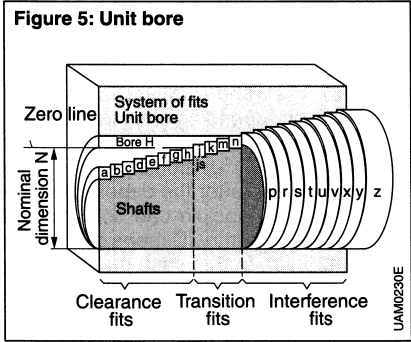
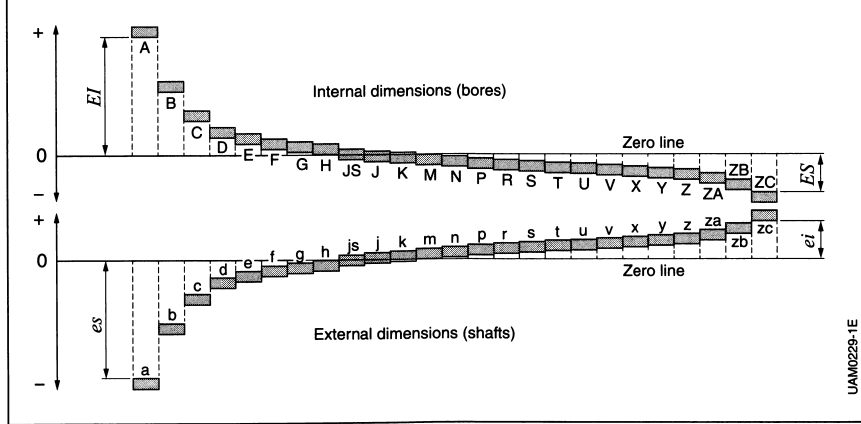


Figure 4: Position of the tolerance intervals

EI Lower limit dimension for internal dimensions, *ES* Upper limit dimension for internal dimensions, *ei* Lower limit dimension for external dimensions, *es* Upper limit dimension for external dimensions.



in many drawings. The indication is provided either in or near the title block. Four tolerance classes (f, m, c, v) are available for dimensional tolerances, while only three tolerance classes (H, K, L) are available for geometrical tolerances. The following entry is required for the tolerance class "middle" (m): ISO 2768–mH.

Surface parameters

The surface texture can dramatically influence the function of a product. Thus, the surfaces of guides and bearings should be as smooth as possible to limit wear and friction losses. However, smooth surfaces are a disadvantage to frictional joints (e.g., driving fits) since these reduce the coefficient of friction and hence also the transmittable force.

Profile meters record the profiles of the real surfaces and evaluate them according to the drawing indications. According to ISO 4287 [10] and ISO 1302 [11] the

following surface parameters can be indicated and evaluated:

- primary profile, unfiltered (parameter P),
- waviness profile, without roughness profile (parameter W),
- roughness profile, without waviness profile (parameter R).

Primarily roughness profiles are used in practice. The parameters most commonly used in drawings are R_z and R_a (Figure 7).

Average peak-to-valley height R_z

The average peak-to-valley height R_z is the mean value of five R_z values from five adjacent individual measured lengths l_i .

$$R_z = \frac{1}{5} \sum_{i=1}^{i=5} R_{z_i}$$

The greatest R_z value from the five individual measured lengths is designated R_{z1max} . R_t is the distance between the highest peak and the lowest valley of the profile of the overall measured length l_n . These two parameters can also be specified to describe the surface.

Arithmetic roughness average R_a

The roughness average R_a is the arithmetical mean of the absolute amounts of the deviations from the surface center line over the measured length l_n .

$$R_a = \frac{1}{l_n} \int_0^{l_n} |Z(x)| dx$$

R_a was the standard characteristic value for surface texture until the 1990s. But since R_a reacts with a lack of sensitivity in respect of peaks and furrows, its informative value is minimal. In other words, R_a

Figure 6: Specification of parallelism

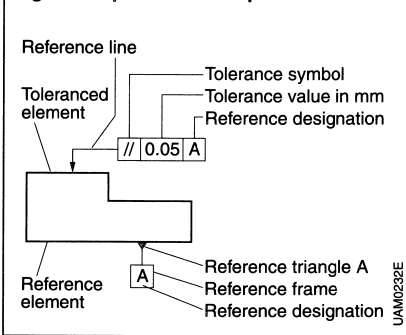


Table 1: Symbols for form and positional tolerances acc. to ISO 1101

Form tolerances without reference		Positional tolerances with reference					
		Direction		Location		Path	
—	Straightness	//	Parallelism	⊕	Position	/	Path
▭	Flatness	⊥	Perpendicularity	◎	Concentricity	↗	Overall path
○	Roundness	∠	Slope	◎	Coaxiality		
⊘	Cylindricity			≡	Symmetry		
⌒	Line form			⌒	Position of line		
⌒	Area form			⌒	Position of an area		

filters very strongly so that for example a single scratch is not evaluated. This is the reason why R_a is used increasingly less and R_z increasingly more for this purpose.

References

- [1] DIN EN ISO 8015: Geometrical product specifications (GPS) – Basic principles – Concepts, principles and rules (ISO 8015:2011); German version EN ISO 8015:2011.
- [2] DIN 7167: Relationship between tolerances of size, form and parallelism; envelope requirement without individual indication on the drawing.
- [3] ISO 14405: Geometrical product specifications (GPS) – Dimensional tolerancing.
- Part 1: Linear sizes.
- Part 2: Dimensions other than linear sizes.
- Part 3: Angular sizes.
- [4] ISO 286: Geometrical product specifications (GPS) – ISO code system for tolerances on linear sizes.
- [5] ISO 1101: Geometrical product specifications (GPS) – Geometrical tolerancing

– Tolerances of form, orientation, location and run-out.

[6] DIN ISO 2768-1: General tolerances; tolerances for linear and angular dimensions without individual tolerance indications.

[7] DIN ISO 2768-2: General tolerances; geometrical tolerances for features without individual tolerance indications.

[8] ISO 8062: Geometrical product specifications (GPS) – Dimensional and geometrical tolerances for molded parts.

Part 1: Terms.

Part 3: General dimensional and geometrical tolerances and machining allowances for castings.

[9] DIN 16742: Plastics molded parts – Tolerances and acceptance conditions.

[10] ISO 4287: Geometrical Product Specifications (GPS) – Surface texture: Profile method – Terms, definitions and surface texture parameters.

[11] ISO 1302: Geometrical Product Specifications (GPS) – Indication of surface texture in technical product documentation.

Figure 7: Surface parameters

a) Arithmetic roughness average R_a .

b) Average peak-to-valley height R_z .

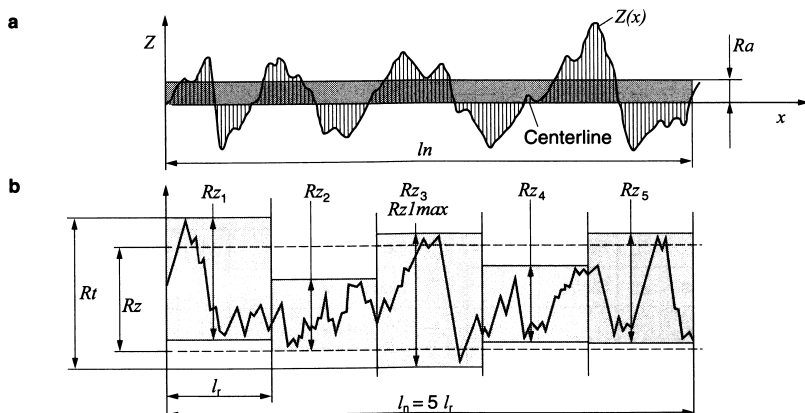
R_t Distance between highest peak and lowest valley of the profile of the overall measured length l_n .

l_i Individual measured length, l_n Overall measured length,

$R_{z1} \dots R_{z5}$ Average peak-to-valley height of the individual measured lengths,

R_{z1max} Greatest R_z value from the five individual measured lengths,

Z Absolute amount of the deviations from the surface center line.



Axles and shafts

Function and application

The purpose of axles and shafts is to absorb forces and torques and transmit them to the casing. But they differ with regard to their function. Thus, an axle never transmits a torque, but a shaft always transmits a torque. While an axle can be stationary or rotate, a shaft is always rotatory. While most shafts rotate (e.g., crankshaft), they can however also perform smaller angular movements (e.g., slewing drives).

Axles

Axle serve exclusively to accommodate wheels, rollers, drums, etc. Rotating machine parts are seated on stationary axles. A rotating axle rotates in bearings about itself, while for example a wheel is permanently mounted on it. Figure 1a shows a fixed axle customary in vehicle manufacturing, while Figure 1b features a rotating axle for a railroad car. The purpose of axles is therefore confined to bearing and supporting. In practice, however, this definition is not always applied consistently, as the example of final drive shows (see Final drive).

Shafts

Shafts must not only accommodate elements and support forces, but also transmit power. Power P is calculated with torque M_t and angular frequency ω or rotational speed n as

$$P = M_t \omega = M_t \cdot 2\pi n.$$

Machine parts such as gears, belt pulleys, clutch hubs and many more are used for input and output. Shafts can be found in every rotating drive system, e.g., the crankshaft in the engine (see Crankshaft) or a shaft in the transmission (Figure 2).

Dimensioning

It is possible to observe in axles and shaft three failure causes, according to which they must then be designed and configured:

- exceeding the load capacity,
- impermissible deformation,
- operation in the resonant range (dynamic response).

Figure 1: Axles

- a) Fixed vehicle axle,
b) Rotating railroad axle.
1 Axle, 2 Wheel, 3 Bearing.
 F_{Whl} Wheel force, F_{Spring} Spring force,
 F_{Brgg} Bearing force.

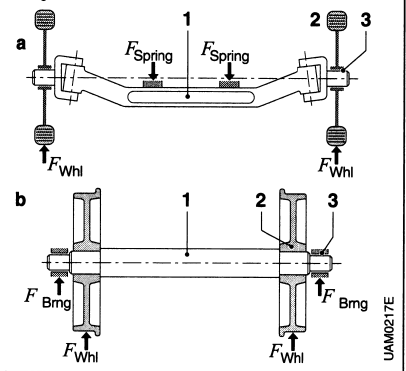
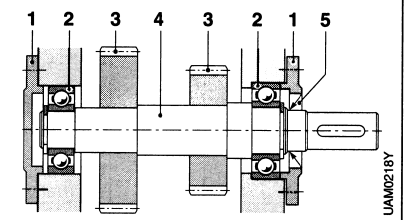


Figure 2: Transmission shaft with locating/ floating bearing arrangement

- 1 Bearing cap, 2 Roller bearing,
3 Gears, 4 Shaft,
5 Shaft seal.



Load capacity

Axles and shafts fail when the existing stresses are greater than the sustainable stresses. The load cases are important as well as the magnitude of the stresses. In other words, whether the stress involved is static or dynamic.

The essential differences between axles and shafts with regard to stresses are set out in Table 1. It can be seen from this that the bending stress is always alternating for axles and shafts. An exception is the eccentric shaft, where the centrifugal force rotates as a result of imbalance with the shaft and hence gives rise to a static bending stress.

In the case of a stationary axle, the bending stress can, depending on the application, be steady, cyclic or alternating.

Shafts are additionally subjected by the torque to torsional stress. Because bending and torsional stresses occur at the same time, a strength hypothesis is required to calculate the existing stress since these stresses cannot simply be summarized by addition. A precise recalculation based on the state of the art can be performed in accordance with DIN 743 [1]. The results are for the most part good, but the work involved is correspondingly high. With less work the load capacity of a shaft can be roughly calculated with the aid of the yield criterion [2].

Deformation

If the permissible deformation is exceeded, axles and shaft can impair the function even when the load capacity is sufficient. Transverse forces give rise to deflections, to which bearings and gears react with great sensitivity (Figure 3). Both the edge loads on the bearings and the engagement faults at the gears generate noises and the service life is greatly reduced as a result. The maximum deflection can be easily calculated for a constant cross-section. However, axles and shafts usually have complex geometries such that deflection lines are determined by means of FEM calculations.

Shafts are additionally twisted by the torsional moment, which can have a negative impact on the transmission behavior of shaft/hub connections. Therefore the torsion angle φ should not exceed 0.25° per meter length. The two ends of a shaft section with constant shaft cross-section at distance l are twisted by the torque M_t through the angle φ :

$$\varphi = \frac{M_t l}{G I_p} \cdot \frac{180^\circ}{\pi}$$

G is the shear modulus and can be taken for steel with sufficient accuracy as 80 GPa. I_p is the polar planar moment of inertia, for solid shafts with diameter d the following applies:

$$I_p = \frac{\pi}{64} \cdot d^4$$

Dynamic response

Shafts are elastic components which have mass and are additionally subjected to masses such as gears, belt pulleys, etc. Therefore rotating axles and shafts are oscillating systems which are excited by centrifugal forces or by rhythmic force or torque fluctuations to forced oscillations. Resonance occurs when the excitation frequency coincides with the natural frequency. The amplitudes are in this case theoretically of infinite magnitude. In other words, an axle or a shaft can already be destroyed under very small loads if the system has enough time to oscillate in the resonant range.

Table 1: Stresses

	Move- ment	Stress	Load case
Axle	stationary	Bending	steady, cyclic, alternating ¹
	rotating	Bending	alternating
Shaft	rotating	Bending	alternating
		Torsion	steady, cyclic, alternating ¹
Eccentric shaft	rotating	Bending Torsion	steady steady, cyclic, alternating ¹

¹ Depending on the application.

Design

Axles and shafts should be able to be produced cost-effectively. In terms of design, this means providing for the fewest possible machined surfaces, using large tolerances, and permitting large peak-to-valley heights. As well as low manufacturing costs, operational reliability must also be guaranteed. The design requirements result from the functions to be performed and the mountability, the required load capacity, the permissible deformation, and the dynamic response.

Function and mountability

A shaft as pictured in Figure 2 must perform a number of functions: directing torque via shaft/hub connections, supporting forces via bearings, preventing axial displacement of the bearing rings, defining axial positions of the gears, and sealing lubricant. A seal makes other demands on the shaft surface than a bearing seat or a shaft/hub connection. It follows from this that as a rule each function has a separate shaft diameter. Securing the axial position of bearing inner rings and hubs can be performed by means of shaft shoulders and retaining rings. When installing the bearings it is important to ensure that no sealing surfaces are damaged. This necessarily results in a stepped shaft with different diameters.

Load capacity

Shaft shoulders, which are required to ensure the function, give rise to stress concentrations as a result of the notch effect and thereby reduce the load capacity. Notches are also created by parallel keys and grooves for retaining rings. Since rotating axles and shafts are dynamically stressed by rotating bending, they are particularly notch-sensitive. Notches must therefore be avoided in highly stressed zones. If this is not possible, the notch effect must be reduced as much as possible by design measures such as small shaft shoulders and larger radii on the inside edge.

Deformation

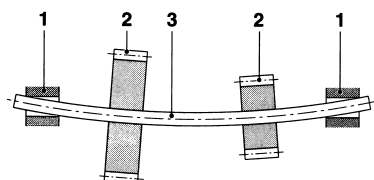
Deformability is essentially dependent on the moment of resistance. Rigid axles and shafts with large moments of resistance are required for small deformations. It is to be noted here that hollow shafts have much better material utilization than solid shafts. Weight savings of up to 50 % are readily possible.

Dynamic response

High torsional and bending stiffness result in high natural frequencies. Therefore rotating axles and shafts are designed to be as stiff as possible so that the resonant speed is sufficiently removed from the operating speed. As with deformations, hollow shafts here too present the advantage that they offer maximum stiffness with minimum mass.

Figure 3: Edge load and misalignment due to deflection

- 1 Bearings,
- 2 Gears,
- 3 Shaft.



UAM0219Y

[1] DIN 743: Calculation of load capacity of shafts and axles.

Part 1: Basic principles.

Part 2: Theoretical stress concentration factors and fatigue notch factors.

Part 3: Strength of materials.

Part 4: Fatigue limit, endurance limit – Equivalently damaging continuous stress.

[2] Haberhauer/Bodenstein: Maschinenelemente. 18th Edition, Verlag Springer Vieweg, 2017.

Springs

Basic principles

Functions

All elastic components to which forces are applied are spring elements. However, springs in the narrower sense mean only those elastic elements that can absorb, store, and release work over a relatively long distance. The stored energy can also be used to maintain a force. The most important applications of industrial springs are:

- Absorbing and damping shocks (shock absorbers and vibration dampers),
- Storing potential energy (spring motors),
- Applying a force (coupling springs),
- Vibrating systems (vibrating table),
- Force measurement (spring balance).

Spring characteristic

The spring characteristic shows the behavior of a spring or spring system. This means the dependency of spring force F or spring torque M_t on deformation. Metal springs have linear characteristics (Hooke's law), rubber springs have progressive characteristics, and disc springs have degressive characteristics (Fig-

ure 1a). The gradient of the characteristic is called the spring rate R .

For translational motion: $R = \frac{dF}{ds}$.

For rotary motion: $R_t = \frac{dM_t}{d\alpha}$.

Spring duty

For frictionless springs under stress, the area under the characteristic represents the absorbed or released work (Figure 1b):

$$W = \int F \, ds .$$

Spring damping

If friction occurs, the prevailing force when the spring is loaded is greater than when the load is removed. The area enclosed by the two characteristics represents frictional work W_R , and is thus a measure of the damping rate (Figure 1b):

$$\psi = \frac{W_R}{W} .$$

Damping due to internal friction can be very high with rubber springs ($0.5 < \psi < 3$). With metal springs, however, it is rather low ($0 < \psi < 0.4$). This means that metal springs have a notable damping rate that is only achievable by means of external friction, e.g. as occurs in layers of leaf and disc springs.

Table 1: Symbols and units

Quantity	Unit
b Width of spring leaf	mm
d Wire diameter	mm
D Mean coil diameter	mm
E Modulus of elasticity	MPa
F Spring force	N
G Shear modulus	MPa
h Height of spring leaf	mm
h_0 Spring deflection (disc spring)	mm
i Number of leaves (leaf spring)	–
i' Number of leaves which continue to ends	–
k Stress coefficient	–
L_c Block length (solid length)	mm
l_i Active length	mm
M_b Bending moment (spring torque)	Nm
M_t Torsional moment (spring torque)	Nm

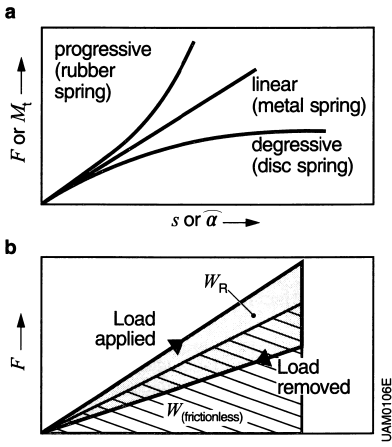
Quantity	Unit
n Number of active coils	–
n_t Total number of coils	–
R Spring rate (spring constant)	N/mm
R_t Torsional spring rate	Nm/rad
s Spring deflection	mm
S_a Total of minimum distances	mm
t Thickness (disc spring)	mm
W Spring duty	J
W_R Frictional work	J
α Twist angle	rad
σ_a Permissible variable stress	MPa
σ_b Bending stress	MPa
σ_m Mean stress	MPa
τ_t Torsional stress	MPa
ψ Damping	–

Spring combinations

A very wide variety of spring characteristics can be achieved by combining several springs (Figure 2). In principle, springs can be combined in parallel or in series. A combination of parallel and series springs is also possible.

Figure 1: Spring characteristics and spring duty

- a) Spring characteristics of different springs,
b) Spring duty.
 s Spring deflection, α Twist angle,
 F Spring force, M_t Spring torque,
 W Spring duty,
 W_R Frictional work.



Parallel combinations

If springs are arranged in parallel, the external load is distributed proportionally between the individual springs. However, the spring deflection (s) is equal for all springs. The spring rate of the spring system is the sum of individual spring rates:

$$R_{\text{total}} = R_1 + R_2 + R_3 + \dots + R_n$$

Accordingly, spring systems comprising parallel springs are harder than individual springs.

Series combinations

With series springs, the total external load acts on each individual spring. However, spring travel for each spring is different depending on the individual spring rates, and are added. The following applies to the resulting spring rate of the overall system:

$$\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n}$$

Spring systems consisting of series springs are softer than the softest individual springs.

Figure 2: Spring combinations

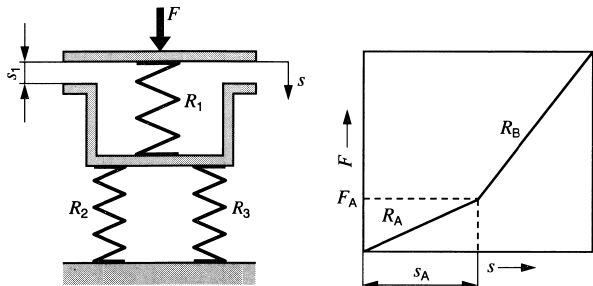
Parallel combination of spring 2 and spring 3 in series with spring 1.

$$R_B = R_2 + R_3$$

$$\frac{1}{R_A} = \frac{1}{R_1} + \frac{1}{(R_2 + R_3)}$$

$$F_A = s_1 R_1$$

$$s_A = s_1 + \frac{F_A}{R_B}$$



Metal springs

Normally, metal springs are classified according to their stresses (Table 2). It should be noted that a loss of power occurs under permanent stress and at higher temperatures; this loss of power is called relaxation. This loss of power, which increases with rising temperature and load duration can be taken from ma-

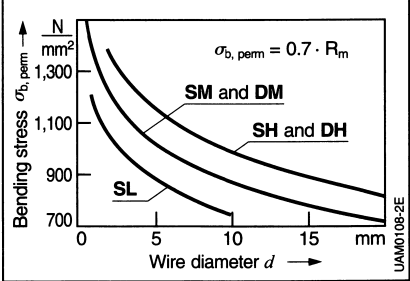
Table 2: Stress on springs

Stress	Construction
Tensile, compressive stress	Tensile test bar, ring spring
Bending stress	Leaf spring, torsion spring, spiral spring, disc spring
Torsional stress	Torsion-bar spring, helical spring

Table 3: Permissible bending stresses for leaf springs

Steel bands	Static stress	Dynamic stress
	$\sigma_{b, perm}$	$\sigma_{b, perm} = \sigma_m \pm \sigma_A$
Hot-rolled	960 MPa	
Cold-rolled, hardened, and draw-tempered	1000 MPa	
Individual leaves ground		500 ± 320 MPa
Individual leaves with rolling skin		500 ± 100 MPa
Layered leaves with rolling skin		500 ± 80 MPa

Figure 3: Permissible bending stresses R_m Tensile strength.



terial-dependent relaxation diagrams for compression springs in DIN EN 13906-1 [1]. From 120 °C, relaxation is no longer negligible. In unalloyed spring steels a loss of power can already occur from 40 °C.

Springs subjected to tensile and compression stress

Due to their high spring stiffness (high spring rate), metal tensile and compression test bars are suitable only for very few special applications.

Springs subjected to bending stress

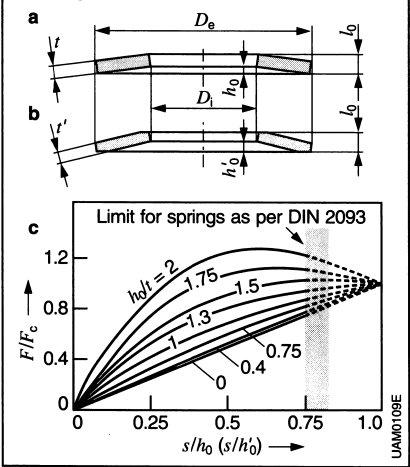
The calculation equations for leaf, coil and spiral springs subjected to bending stress are set out in Table 4.

Leaf springs

A simple leaf spring is used as a compression or guide spring. Layered leaf springs are used for suspension and wheel control in vehicles. They are usually made of spring steel in accordance with DIN EN 10089 [2] (hot-rolled) and DIN EN 10132 [3] (cold-rolled strip). The draft design may assume the permissible bending stresses specified in Table 3.

Figure 4: Disc springs according to DIN 2093

- a) Without contact bearing surfaces,
- b) With contact bearing surfaces,
- c) Calculated spring characteristic of a disc spring to DIN 2092.



Torsion and spiral springs

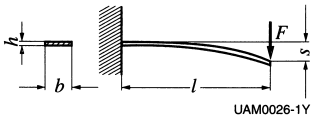
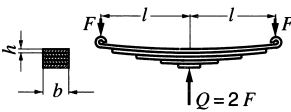
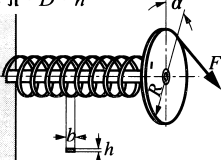
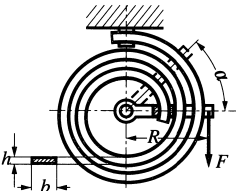
During an excursion of the torsion or spiral spring (twist angle α) aligning torques are generated about the rotational axis. Due to the clamping conditions, the bending stresses in the angle are almost uniform. The calculation equations for circular and rectangular sections, which apply to both torsion and spiral springs, are set out in Table 4. For the frequently used spring-steel grades (SL, SM or DM and SH or DH) it is possible, for design purposes in the case of static and quasi-static stress, when the increase in stress as a result

of the wire curvature is disregarded, to calculate with the permissible bending stresses in accordance with Figure 3.

Disc springs

The cone-bowl-shaped disc springs (Figure 4) are primarily subjected to bending stresses. A wide variety of applications results from the large number of possible parallel and series combinations. Disc springs are mainly used where spring force and travel must be absorbed within confined spaces. They are used among other things in regular and

Table 4: Leaf spring, torsion spring, and spiral spring

Type	Spring force	Spring deflection
Simple straight leaf spring (constant cross-section)  UAM0026-1Y	$F_{\max} = \frac{b \cdot h^2}{6 \cdot l} \cdot \sigma_{b, \text{perm}}$	$s = \frac{4 \cdot F \cdot l^3}{E \cdot b \cdot h^3}$
	Spring rate $R = \frac{F}{s} = \frac{E \cdot b \cdot h^3}{4 \cdot l^3}$	Spring duty $W_{\max} = \frac{\sigma_{b, \text{perm}}^2}{18 \cdot E} \cdot b \cdot h \cdot l$
Layered leaf spring  UAM0028-1Y	Spring force, spring torque $F_{\max} = \frac{i \cdot b \cdot h^2}{6 \cdot l} \cdot \sigma_{b, \text{perm}}$	Spring deflection $s = \frac{12 \cdot F \cdot l^3}{(2 \cdot i + i') \cdot E \cdot b \cdot h^3}$
	Spring rate $R = \frac{(2i + i') \cdot E \cdot h^3 \cdot b}{12 \cdot l^3}$	Spring duty $W_{\max} = \frac{\sigma_{b, \text{perm}}^2}{3 \cdot E} \cdot \frac{l^2}{(2i + i')} \cdot b \cdot h \cdot l$
Torsion spring (leg spring) $l_t = \pi \cdot D \cdot n$  UAM0029-1Y	Spring force, spring torque $M_{t, \max} = \frac{\pi \cdot d^3}{32} \cdot \sigma_{b, \text{perm}}$	Spring deflection $\hat{\alpha} = \frac{64 \cdot M_t \cdot l}{E \cdot \pi \cdot d^4}$
	Spring rate $R = \frac{M_t}{\hat{\alpha}} = \frac{E \cdot \pi \cdot d^4}{64 \cdot l}$	Spring duty $W_{\max} = \frac{\sigma_{b, \text{perm}}^2}{32 \cdot E} \cdot \pi \cdot d^2 \cdot l$
Spiral spring $l_s = 2 \cdot \pi \cdot n \cdot [r_0 + 0.5 \cdot n \cdot (h + \delta_r)]$  UAM0030-1Y	Spring force, spring torque $M_{t, \max} = \frac{b \cdot h^2}{6} \cdot \sigma_{b, \text{perm}}$	Spring deflection $\hat{\alpha} = \frac{12 \cdot M_t \cdot l}{E \cdot b \cdot h^3}$
	Spring rate $R = \frac{M_t}{\hat{\alpha}} = \frac{E \cdot b \cdot h^3}{12 \cdot l}$	Spring duty $W_{\max} = \frac{\sigma_{b, \text{perm}}^2}{6 \cdot E} \cdot b \cdot h \cdot l$

overload clutches and to pre-tension rolling bearings. Because friction occurs between the individual disc springs in spring assemblies layered in the same direction, they are also suitable for damping vibrations and shocks.

At $h_0/t > 0.4$, the nonlinearities of the springs are no longer negligible. Spring forces, spring travel, and spring rates can be calculated with sufficient precision according to DIN 2092 [4], or they can be taken from the manufacturer's data.

For disc springs subjected to static stress ($< 10^4$ stress reversals), fatigue-strength calculation is not required if the maximum spring force at $s = 0.75 h_0$ is not exceeded. For dynamically stressed disc springs the maximum permissible tensile stress and

the permissible stress range can be taken from DIN 2093 [5].

Springs subjected to torsional stress

The calculation equations for torsion-bar and helical springs subjected to torsional stress are set out in Table 5.

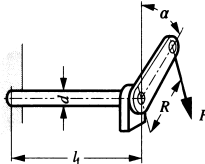
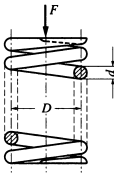
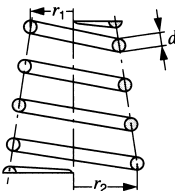
Torsion-bar springs

Circular sections are usually selected for torsion bars. They have a very high volume utilization factor, which means that they can absorb a lot of energy, but they occupy little space.

Helical springs

Cylindrical helical springs are manufactured as compression and extension

Table 5: Torsion-bar spring and helical springs

Type	Spring force, spring torque	Spring deflection
Torsion-bar spring Circular section (DIN 2091, [6]) 	$M_{t, \max} = \frac{\pi \cdot d^3}{16} \cdot \tau_{t, \text{perm}}$	$\hat{\alpha} = \frac{32 \cdot M_t \cdot l_t}{G \cdot \pi \cdot d^4}$
Spring rate	Spring duty	
$R = \frac{M_t}{\hat{\alpha}} = \frac{G \cdot \pi \cdot d^4}{32 \cdot l_t}$	$W_{\max} = \frac{\tau_{t, \text{perm}}^2}{16 \cdot G} \cdot \pi \cdot d^2 \cdot l_t$	
Cylindrical helical springs Circular section (DIN EN 13906, [1]) Druckfeder  Zugfeder 	$F_{\max} = \frac{\pi \cdot d^3}{8 \cdot k \cdot D} \cdot \tau_{t, \text{perm}}$	$s = \frac{8 \cdot D^3 \cdot n \cdot F}{G \cdot d^4}$
Spring rate	Spring duty	
$R = \frac{G \cdot d^4}{8 \cdot D^3 \cdot n}$	$W_{\max} = \frac{\tau_{t, \text{perm}}^2}{16 \cdot G} \cdot d^2 \cdot D \cdot \pi^2 \cdot n$	
Tapered helical springs with circular section 	$F_{\max} = \frac{\pi \cdot d^3}{16 \cdot k \cdot r_2} \cdot \tau_{t, \text{perm}}$	$s = \frac{16 \cdot (r_1 + r_2) \cdot (r_1^2 + r_2^2) \cdot n \cdot F}{G \cdot d^4}$
Spring rate	Spring duty	
$R = \frac{G \cdot d^4}{16 (r_1 + r_2) \cdot (r_1^2 + r_2^2) \cdot n}$	$W_{\max} = \frac{\tau_{t, \text{perm}}^2}{32 \cdot G} \cdot \frac{d^2 (r_1 + r_2) \cdot (r_1^2 + r_2^2) \cdot \pi \cdot n}{r_2^2}$	

springs. The calculation equations are identical for both types. Compression springs in conical form can optimize the use of space if the individual coils can be pushed into each other.

Compression springs

To enable the force to be introduced as centrally as possible in the case of a compression spring, the spring ends are created and surface-ground vertically to the spring axis. Cold-formed springs consist of at least two active coils ($n \geq 2$) and two created, non-active end coils. In the case of hot-formed springs n should be ≥ 3 and as non-active end coils only $\frac{3}{4}$ coils are created in each case. This produces the total number of active coils as indicated in Table 6. Furthermore, the spring ends are arranged with an offset of 180° to each other so as to produce a total coil number of $n_t = 2.5; 3.5; 4.5; 5.5; 6.5$, etc.

In order to avoid plastic deformation in the spring, the active coils may not touch each other under maximum load. In other words, a compression spring must never be compressed to the block (solid) length

Figure 6: Permissible torsional stresses for helical springs with static stress

- a) Cold-formed from patented drawn spring-steel wires (SL, SM, DM, SH and DH) and valve spring-steel wire (VDC) acc. to DIN EN 10270-2,
b) Hot-formed from spring steels acc. to DIN EN 10089.
 R_m Tensile strength.

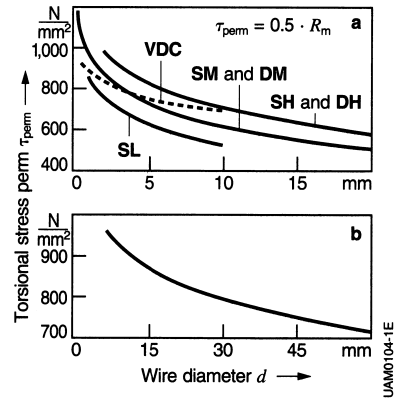
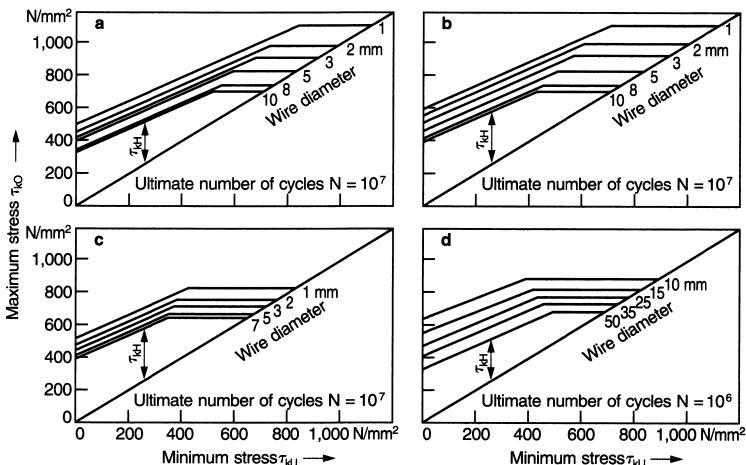


Figure 5: Fatigue-limit diagrams for helical compression springs

- a) For cold-formed springs made from spring-steel wires SH and DH (not shot-blasted),
b) For cold-formed springs made from spring-steel wires SH and DH (shot-blasted),
c) For cold-formed springs made from valve spring-steel wire (VDC),
d) For hot-formed springs.



L_c . The sum of the minimum distances between the individual coils is indicated with S_a and can be calculated in accordance with Table 6. Under dynamic load S_a must be doubled for hot-formed springs, whereas a factor of 1.5 should be used for cold-formed springs.

The effect of the wire curvature is taken into account by the stress coefficient k (Table 7). In the case of static stress, this effect can be disregarded, i.e. $k = 1$ is then set. The following applies to the stress range prevailing in the case of dynamic stress:

$$\tau_{kh} = k \frac{8 D}{\pi \cdot d^3} \cdot (F_2 - F_1) \leq \tau_{kH} .$$

Extension springs

Extension springs are formed either with loops or with rolled-in or screwed-in end pieces. Since service life is determined primarily by the loops, it is impossible to give general fatigue limit values. Cold-formed extension springs hardened and tempered after drawing can be manufactured with an internal preload. This allows significantly higher stresses.

References

[1] DIN EN 13906: Cylindrical helical springs made from round wire and bar – Calculation and design. Part 1: Compression springs.
Part 2: Extension springs.
Part 3: Torsion springs.

[2] DIN EN 10089: Hot rolled steels for quenched and tempered springs – Technical delivery conditions.
[3] DIN EN 10132: Cold-rolled narrow steel strip for heat-treatment – Technical delivery conditions.
Part 1: General.
Part 2: Case hardening steels.
Part 3: Steels for quenching and tempering.
Part 4: Spring steels and other applications.
[4] DIN 2092: Disc springs – Calculation.
[5] DIN 2093: Disc springs – Quality requirements – Dimensions.
[6] DIN 2091: Circular section torsion bar springs; calculation and design.
[7] DIN-Taschenbuch 29: Federn 1. Berechnungs- und Konstruktionsgrundlagen, Qualitätsanforderungen, Bestellangaben, Begriffe, Formelzeichen und Darstellungen. Beuth-Verlag 2015.
[8] DIN-Taschenbuch 349: Federn 2. A compilation of the current material standards and definitions for semi-finished products. Beuth-Verlag 2012.
[9] H. Haberhauer; F. Bodenstein: Maschinenelemente. 18th Edition, Verlag Springer Vieweg, 2017.
[10] F. Fischer; H. Vondracek: Warm geförmte Federn – Konstruktion und Fertigung. Hohenlimburg – Hoesch AG, 1987.
[11] M. Meissner; H.-J. Schorcht; U. Kletz: Metallfedern – Grundlagen, Werkstoffe, Berechnung, Gestaltung und Rechnereinsatz. 3rd Ed., Verlag Springer Vieweg, 2015.

Table 6: Helical springs

	Total number of coils	Block length (solid length)	Sum of minimum distances
Cold-formed	$n_1 = n + 2$	$L_c \leq n_1 \cdot d$	$S_a = (0.0015 \cdot D^2/d + 0.1 \cdot d) \cdot n$
Hot-formed	$n_1 = n + 1.5$	$L_c \leq (n_1 - 0.3) \cdot d$	$S_a = 0.02 \cdot (D + d) \cdot n$

Table 7: k-factor

D/d	3	4	6	8	10	14	20
k	1.55	1.38	1.24	1.17	1.13	1.10	1.07

Friction bearings

Features

The task of bearings is to guide machine parts and support forces between surfaces that are moving relative to each other. In order to minimize the wear and power loss, the friction that occurs in the process must be kept as low as possible.

Friction (or plain or sliding) bearings are made of metallic (e.g. sintered metals) and non-metallic materials. A distinction is made between radial and axial (thrust) bearings based on the direction of the external load. Another distinguishing feature is lubrication. Hence, friction bearings can be either hydrostatically or hydrodynamically lubricated, or self-lubricating. A complete separation of the sliding surfaces is the optimal state. With full lubrication there is pure fluid friction with very low coefficients of friction. The lubricant layer has a shock-, vibration- and noise-absorbing effect.

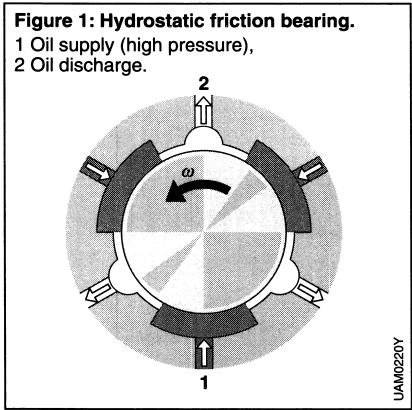
Table 1: Symbols and units
(DIN 31652 [1])

Description	Symbol	Unit
Axial bearing length	B	m
Inside bearing diameter (nominal diameter)	D	m
Shaft diameter (nominal diameter)	d	m
Eccentricity (displacement between shaft and bearing centers)	e	m
Bearing force (load)	F	N
Minimum lubricating film thickness	$h_{\min}$	m
Local lubricating film pressure	p	$\text{Pa} = \text{N/m}^2$
Specific bearing load $= F/(B D)$	$\bar{p}$	Pa
Bearing clearance $= D - d$	s	m
Sommerfeld number	So	–
Sliding velocity	v	m/s
Relative eccentricity $= 2e/s$	ε	–
Effective dynamic viscosity of lubricant	η_{eff}	$\text{Pa} \cdot \text{s}$
Relative bearing clearance $= s/D$	ψ	–
Hydrodynamically effective angular velocity	ω_{eff}	s^{-1}

Hydrostatic friction bearings

In the case of hydrostatic bearings the lubricant is forced by an external pump at high pressure between the sliding surfaces. At any point in time there is a complete separation of the sliding surfaces on account of the thin lubricating film (Figure 1). Friction losses only occur on account of the shear forces in the lubricant. These forces are proportional to the speed at which the bearing surfaces move against each other. At low relative speeds hydrostatic friction bearings consequently operation with virtually zero friction. There is no stick-slip effect when starting and after-running due to the absence of slip resistance.

Hydrostatic friction bearings are therefore wear-and-tear-resistant and consequently have a very long service life. The drawbacks of these friction bearings are the high costs and the large installation space required. For these reasons they are used primarily in heavy machine construction.



Hydrodynamic friction bearings

Application

Most of the hydrodynamic friction bearings used in motor-vehicle engines are plain bearings (Figure 2) for holding the crankshaft drive, including the camshafts. They are usually designed as bearing shells with special clearance (oval clearance acc. to Figure 3).

Operating principle

In the case of hydrodynamic friction bearings the pressure is generated automatically. The following preconditions are required for this purpose:

- relative movement between the sliding surfaces,
- wedge-shaped gap,
- lubricant must adhere to the sliding surfaces,
- sufficient lubricant supply.

A friction bearing requires a bearing clearance in order to build up a lubricating gap that is capable of bearing. In the stationary state the shaft touches the bearing shell and consequently there is solid-body friction. The eccentric displacement of the shaft creates a wedge-shaped lubricating

gap (Figure 3). As the rotational speed increases a lubricating film that is capable of bearing is built up as a result of the lubricant that adheres to the sliding surfaces being pumped into the wedge gap. Until the pressure has built up fully in the lubricating gap, i.e. until the sliding surfaces are completely separated, the shaft still partly touches the bearing shell. During this period the bearing is in the mixed-friction range. There is pure fluid friction only when the operating rpm exceeds the transition rpm.

Stribeck curve

The friction states in the hydrodynamic friction bearing were analyzed by Richard Stribeck. The curve named for him shows the dependence of friction on shaft speed (Figure 4, [2]).

Sommerfeld number

Depending on load and rotational speed, the eccentricity of the shafts is set within the system in such a way that the integral of the lubricant pressure of the external bearing force maintains equilibrium. The hydrodynamic pressure distribution in a convergent bearing gap is determined from the solution of Reynolds' differential equation. The integration of the pressure

Figure 2: Pressure distribution in plain bearing

- e Eccentricity,
- F Bearing force,
- $h_{\min}$ Smallest lubricating gap,
- d Shaft diameter,
- D Bearing inside diameter,
- p Pressure distribution,
- ω Angular velocity,
- S Shell center point (bearing),
- W Shaft center point.

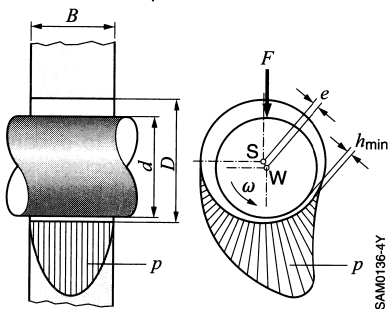
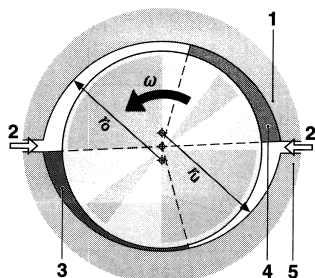


Figure 3: Two-face bearing shell with oval clearance

- 1 Upper bearing shell,
- 2 Oil supply,
- 3 Lower lubricating wedge,
- 4 Upper lubricating wedge,
- 5 Lower bearing shell.
- r_o Radius, upper bearing shell,
- r_u Radius, lower bearing shell,
- ω Angular velocity.



distribution produces the load capacity of the lubricating film and is expressed in the dimensionless Sommerfeld number So [1] (for quantities, see Table 1):

$$So = \frac{F \psi^2}{(D B \eta_{eff} \omega_{eff})}$$

For stable operation $So > 1$ must be the case, bearings subjected to high load should be designed with $So > 3$. As the Sommerfeld number increases, so the relative eccentricity ϵ increases and the minimum lubricating-gap thickness h_{min} decreases. The following applies:

$$h_{min} = \frac{(D - d)}{2} - e = 0.5 D \psi (1 - \epsilon)$$

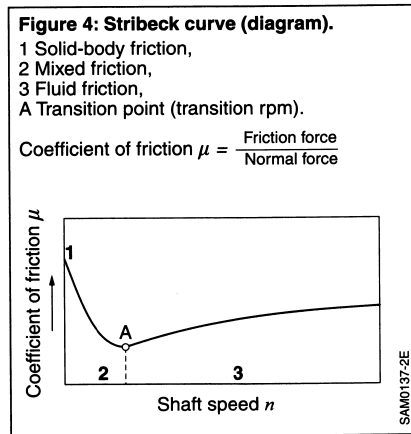
The relative eccentricity is then:

$$\epsilon = \frac{2e}{(D - d)}$$

With the aid of the Sommerfeld number it is possible, in accordance with DIN 31652 [1] and VDI 2204 [2], to calculate the coefficient of friction μ in the bearing and thus also the friction loss and the thermal stress.

Table 2: Coefficients of friction for different types of friction

Type of friction	Coefficient of friction μ
Solid-body friction	0.1 ... > 1
Mixed friction	0.01 to 0.1
Fluid friction	0.001 to 0.01



Requirements

The coefficients of friction given in Table 2 are approximate values, and are intended only for comparison of the different operating states. As hydrodynamic bearings also operate with mixed friction some of the time, must be able to withstand a certain amount of contamination without loss of function, and are also subjected to high dynamic and thermal stress (particularly in piston engines), the bearing material must meet a number of requirements, some of which are mutually exclusive.

- Conformability: This is the property of a bearing material to adapt to necessary modifications of shape by local deformation without permanent damage.
- Embeddability: This is the ability to absorb dirt particles into the bearing surface without negative consequences.
- Resistance to wear: This prevents the separation of small particles in the event of mechanical load in the mixed-friction area.
- Run-in performance: This is the interaction of conformability, embeddability, and resistance to wear.
- Resistance to seizing: This prevents partial welding of the sliding surfaces under high compressive load.
- Mechanical loadability: This prevents plastic deformation under high compressive loads.
- Fatigue strength: This describes the slowly progressing material fatigue which under alternating load can result in failure even in the case of small bearing forces.

If a bearing (e.g. piston-pin bushing) is simultaneously subjected to high loads and low sliding velocities, high fatigue strength and wear resistance should take precedence over resistance to seizing. The bearing materials used in such cases are hard bronzes or special brass alloys.

Since they are subjected to dynamic loads with high sliding velocities, connecting-rod and crankshaft bearings in internal combustion engines must fulfill a number of different requirements. In these applications, multilayer bearings (Figure 5), above all trimetal bearings, have proved themselves in practice.

The service life of friction bearings in the crankshaft drive can be further increased through the use of special solutions such as the sputter bearing (Figure 6). It contains a highly wear-resistant AlSn running layer (sputter layer) which is applied in the PVD procedure (physical vapor deposition) to the bearing material underneath. Grooved sliding bearings (Figure 7) are also used in internal-combustion engines subjected to high load (e.g. turbo-

charged diesel engines). In this variant fine grooves, which are filled with a soft filler (e.g. PbSnCu), are incorporated in the circumferential direction of the bearing surface. The bearing surface therefore exhibits soft and hard areas next to each other.

Bearing materials

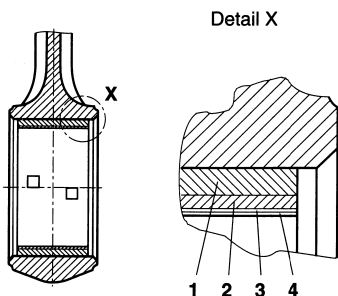
Lead, tin, copper, and aluminum alloys are used as bearing materials. Values for the permissible specific bearing load are given in Table 3.

Lead and tin bearing metals, formerly called babbitt metals, are ideally suitable for higher sliding velocities and are characterized by good run-in and emergency-running properties. Tin bronzes

Figure 5: Multilayer bearing

(Design of a trimetal bearing).

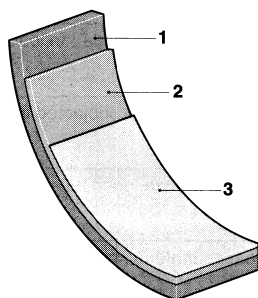
- 1 Steel backing shell, 2 Bearing metal,
- 3 Diffusion barrier (e.g. 1 – 2 μm nickel),
- 4 Penetration coating (approx. 20 μm , electroplated SnCu layer or anti-friction paint).



UAM0020Y

Figure 6: Section through a sputter bearing (lead-free)

- 1 Steel backing,
- 2 Intermediate layer (brass or bronze),
- 3 Running layer (e.g. AlSn 20 Cu).



SAM0138Y

Table 3: Empirical values for permissible specific bearing load (acc. to DIN 31652-3)

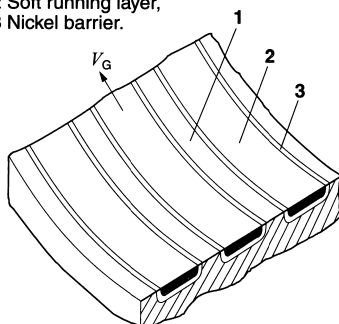
Bearing materials	Permissible specific bearing load $\bar{p}_{lim}$ [MPa]
Pb and Sn alloys (babbitt metals)	5 to 15
Bronze, tin base	7 to 25
Bronze, lead base	7 to 20
AlSn alloys	7 to 18
AlZn alloys	7 to 20
Maximum values only apply for very low sliding velocities	

Figure 7: Section through a grooved sliding bearing

(Miba patent).

Running surface with very fine grooves in running direction V_G .

- 1 Wear-resistant light alloy,
- 2 Soft running layer,
- 3 Nickel barrier.



UAM0021Y

are suitable for high stresses and are very wear-resistant; run-in and emergency-running properties are however less good. Lead bronzes demonstrate better emergency-running properties with only slightly worse resistance to wear. Aluminum alloys have a higher corrosion resistance than babbitt metal and copper alloys (bronzes).

Standardized materials for friction bearings are listed in ISO 4381 [3], ISO 4382 [4], and ISO 4383 [5]. Some bearing materials contain lead. EU 2016/774 [6] prohibits the use of lead in passenger cars. Lead is currently still permitted in commercial vehicles and in general machine construction.

Self-lubricating bearings made of metal

Sintered bearings

Friction bearings made of sintered metal belong to the category of self-lubricating bearings. They consist of sintered metals which are porous and impregnated with liquid lubricants. For many small motors in motor automotive applications, this type of bearing is a good compromise in terms of precision, installation, freedom from maintenance, service life, and cost. They are used primarily with shaft diameters from 1.5 to 12 mm.

For motor vehicles sintered-iron and sintered-steel bearings are preferred to sintered-bronze bearings since they are cheaper and exhibit a smaller interaction with the lubricant (Table 4). The advantages of sintered-bronze bearings are higher permissible sliding velocities, lower noise, and lower friction coefficients (e.g. for record-players, office equipment, and data systems).

Sintered metals bearing the designation SINT-B (Table 4) have 20 % porosity. Aside from these there are also SINT-A with 25 % and SINT C with 15 % porosity.

The performance of sintered bearings over long periods of service is closely related to the use of optimum lubricants. Mineral oils are used; however, these exhibit insufficient cold-flow properties and moderate resistance to aging. Synthetic oils (e.g. esters, poly- α -olefins) on the other hand have good cold-flow proper-

Table 4: Materials for sintered-metal bearings (acc. to DIN 30910-3 [7])

Material group	Designation Sint- ...	Composition [%]	Remarks
Sintered iron	B 00	< 0.3 C < 1 Cu Rest Fe	Standard material which meets moderate load and noise requirements.
Sintered steel, containing Cu	B 10	< 0.3 C 1...5 Cu Rest Fe	Good resistance to wear, can be subjected to higher loads than pure Fe bearing.
Sintered steel, higher Cu content	B 20	< 0.2 Cu 15...20 Cu Rest Fe	Cheaper than sintered bronze, good noise behavior.
Sintered bronze	B 50	< 0.2 C 9...11 Sn Rest Cu	Standard Cu-Sn based material, good noise behavior.

ties and can be subjected to high thermal loads. Furthermore, the evaporation tendency is low. Synthetic grease oils (oils which have metal soaps) are characterized by low starting friction and low wear. The most important properties of sintered bearings are set out in Table 5.

Metal-ceramic bearings

Metal-ceramic friction bearings consist of material manufactured by powder-metal-lurgy processes; in addition to the metallic matrix, the bearing material also contains finely distributed solid lubricant particles.

Bronze, iron, and nickel are used as the material, graphite or molybdenum disulfide (MoS_2) for example are used as the lubricant. Ceramic bearings are particularly suitable for use under high loads, and are at the same time self-lubricating. However, they are very brittle and therefore sensitive to jolts and impacts.

Metal-ceramic bearings are used in motor vehicles, for example as steering-knuckle bearings.

Table 5: Properties of maintenance-free, self-lubricating bearings

Properties	Sintered bearings oil-impregnated		Polymer bearings		Metal-backed composite bearings Running layer		Synthetic carbons
	Sintered iron	Sintered bronze	Thermoplastic polyamide	Duroplastic polyimide	PTFE + additive	Acetal resin	
Pressure resistance MPa	80...180		70	110	250	250	100...200
Max. sliding velocity m/s	10	20	2	8	2	3	10
Specific load MPa	1...4 (10) ¹		15	50 (at 50 °C) 10 (at 200 °C)	20...50	20...50	50
Permissible operating temperature °C	-60...180 (depends on oil) 200		-130...100	-100...250	-200...280	-40...100	-200...350
Short-term			120	300		130	500
Coefficient of friction without lubrication	with lubrication 0.04...0.2		0.2...0.4 (100 °C) 0.4...0.6 (25 °C)	0.2...0.5 (unfilled) 0.1...0.4 (filled)	0.4...0.2	0.7...0.2 PTFE filled	0.1...0.35
Thermal conductivity W/(m·K)	20...40		0.3	0.4...1	46	2	10...65
Corrosion resistance	less good	good	very good		good	good	very good
Chemical resistance	no		very good		conditional	conditional	good
Max. $p \cdot v$ MPa·m/s	20		0.05	0.2	1.5...2		0.4...1.8
Embeddability of dirt and abraded material	less good		good	good	less good	good	less good

¹ Value in parentheses applies to additional lubrication.

Self-lubricating bearings made of plastic

Different plastics can be used for friction bearings. The various properties of plastic bearings are set out in Table 5.

Solid polymer bearings made of thermoplastics

Features

Friction bearings made of thermoplastics are inexpensive and are suitable for low bearing forces and low operating temperatures. The danger of "seizing" is extremely small.

The most frequently used thermoplastic polymer materials are:

- Polyoximethylene (POM, POMC)
- Polyamide (PA)
- Polyethylene and polybutylene terephthalate (PET, PBT)
- Polyetheretherketone (PEEK)

The tribological and mechanical properties can be varied over a wide range by incorporating lubricants and reinforcements in the thermoplastic base material.

Lubrication additives

- Polytetrafluoroethylene (PTFE)
- Graphite (C)
- Silicone oil, and other liquid lubricants, recently also enclosed in microcapsules

Reinforcement additives

- Glass fibers (GF)
- Carbon fibers (CF)

Application examples

- Windshield-wiper bearings (PA and glass fiber)
- Idle actuators (PEEK + carbon fiber, PTFE and other additives)

Polymer bearings made of duroplastics

These materials have high intrinsic friction and are seldom used as bearing materials in motor vehicles. Duroplastics which are used for friction bearings are:

- Phenolic resins (high friction)
- Epoxy resins (additives of PTFE or C, reinforcement by fibers required on account of their inherent brittleness)
- Polyimides (high thermal and mechanical loadability)

Application example

Polymer bearings made of duroplastics are used in wiper motors as axial startup blocks made of polyimide.

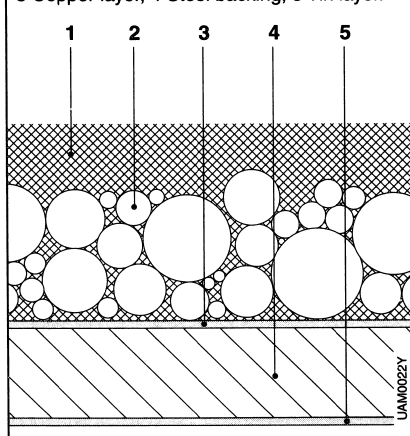
Metal-backed composite bearings

Composite bearings are combinations of polymer materials, fibers, and metals. The respective structure (Figure 8) deliver advantages over pure or filled polymer friction bearings with regard to load capacity, bearing clearance, heat conduction, and installation. They are also more suitable for oscillating motions.

Example of bearing structure

The bearing consists of a steel backing-
ing electroplated with tin; above this a

Figure 8: Section through a self-lubricating composite bearing
1 Polymer liner, 2 Porous bronze layer, 3 Copper layer, 4 Steel backing, 5 Tin layer.



0.2...0.35 mm thick bronze-bead layer with 30...40 % porosity in which a low-friction polymer material is rolled in as a liner is sintered on. The liner consists of

- acetal resin or polyvinylidene fluoride, either impregnated with oil or containing lubricating recesses, or
- PTFE + ZnS or molybdenum disulfide (MoS_2) and graphite as additive.

Metal-backed composite bearings are available in a number of different shapes and compositions. Metal-backed composite bearings with woven PTFE fiber inserts have unusually high loadability and are suitable for use in ball-and-socket joints.

Application examples for motor vehicles

- Piston-rod bearings for suspension struts
- Release-lever bearings for clutch pressure plates
- Brake-shoe bearings in drum brakes
- Ball-and-socket bearings
- Door-hinge bearings
- Bearings for winding shafts for seat belts
- Steering-knuckle bearings
- Gear-pump bearings

Composite bearings with specially modified liners are required for heavy-duty requirements in diesel high-pressure injection pumps. The liner is made of PEEK or PPS with additives (e.g. carbon fibers, ZnS, TiO_2 , and graphite). The particle size is sometimes in the nanometer range.

Carbon-graphite bearings

Carbon-graphite bearings are members of the ceramic bearing family due to their method of manufacture and material properties. Powdered carbons are used as the base material, pitch and synthetic resins are used as binders. It is to be noted that carbon-graphite bearings are very brittle.

Advantages

- Thermal stability up to 350 °C (hard-burnt carbon) and up to 500 °C (electro-graphite)
- Low friction
- Good corrosion resistance
- Good thermal conductivity
- Good thermal shock resistance.

Application examples for carbon-graphite bearings

- Fuel-pump bearings
- Bearings in drying ovens
- Adjustable guide vanes of compressors in exhaust-gas turbochargers

References

- [1] DIN 31652: Plain bearings – Hydrodynamic plain journal bearings under steady-state conditions.
Part 1: Calculation of circular cylindrical bearings.
Part 2: Functions for calculation of circular cylindrical bearings.
Part 3: Permissible operational parameters for calculation of circular cylindrical bearings.
- [2] VDI 2204: Design of plain bearings.
- [3] ISO 4381: Plain bearings – Tin casting alloys for multilayer plain bearings.
- [4] ISO 4382: Plain bearings; copper alloys;
Part 1: Cast copper alloys for solid and multilayer thick-walled plain bearings.
Part 2: Wrought copper alloys for solid plain bearings.
- [5] ISO 4383: Plain bearings – Multilayer materials for thin-walled plain bearings.
- [6] Commission Directive (EU) 2016/774 of 18 May 2016 amending Annex II to Directive 2000/53/EC of the European Parliament and of the Council on end-of-life vehicles.
- [7] DIN 30910: Sintered metal materials; sintered-material specifications (WLB);
Part 1: Explanatory notes for WLB.
Part 2: Sintered metal materials for filters.
Part 3: Sintered metal materials for bearings and structural parts with bearing properties.
Part 4: Sintered metal materials for structural parts.
Part 6: Hot-forged sintered steels for structural parts.

Rolling bearings

Application

Rolling bearings are some of the most important components in machines. Great demands are made on their load capacity and operational reliability. Rolling bearings are widely used in motor vehicles. Passenger cars and commercial vehicles are fitted with a large number of rolling bearings in bearing arrangements for alternator, starter, wheel bearing, transmission, suspension strut, drive shaft, water pump, tension pulley, steering, windshield-wiper motor, fan, and fuel-injection pump.

General principles

Type

Rolling bearings are generally made up of two races (Figure 1, outer race and inner race), a cage, and a rolling-element assembly. The rolling elements guided by the cage roll on the raceways. Balls, cylindrical rollers, needle rollers, tapered rollers, and self-aligning rollers are used as rolling elements. A rolling bearing can be lubricated with grease. It is fitted with sheet-steel cover plates or rubber gaskets to prevent grease discharge and to provide a seal against dirt.

The rolling bearing transfers the outer force from one bearing race to the other via the rolling elements. A distinction is made between radial bearings and axial (thrust) bearings, depending on the main direction of load.

Structural dimensions

The rolling bearing is a ready-to-install machine part. Various ranges of outside diameters and widths are available for a hole diameter. Standardized abbreviation codes are used to identify the diameter and width ranges of a rolling bearing. The designations are recorded in the standard DIN 623-1 [1].

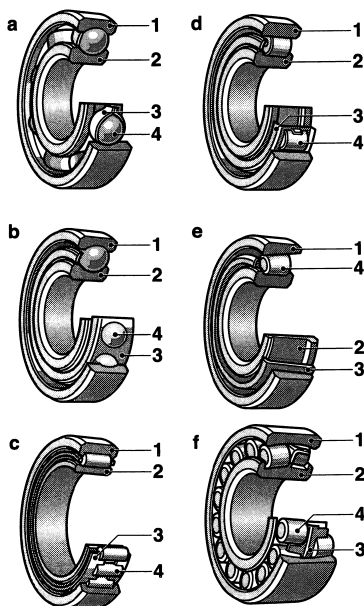
The external dimensions are indicated in the catalogs of rolling-bearing manufacturers.

Tolerances

Dimensional and form tolerances of rolling bearings are standardized in ISO 492 [2] and DIN 620 ([3], [4], [5], [6], [7]) according to precision. Rolling bearings of normal precision, i.e. tolerance class P0 (also called PN), generally satisfy all the demands which are made by mechanical engineering on bearing quality. For more stringent requirements, the standard provides for more precise tolerance classes P6, P5, P4, and P2.

Figure 1: Design of rolling bearings

- a) Deep-groove ball bearing,
 - b) Angular-contact ball bearing,
 - c) Needle bearing,
 - d) Cylindrical rolling bearing,
 - e) Tapered rolling bearing,
 - f) Self-aligning rolling bearing.
- 1 Outer race, 2 Inner race,
3 Cage, 4 Rolling element.



UAM0117Y

Bearing play and bearing clearance

The bearing play of an uninstalled rolling bearing refers to the distance through which the bearing races can be moved against each other. A distance is made between radial play and axial play.

Radial play is defined in the standard DIN 620 Part 4 [6]. The normal radial play category is CN. In accordance with operating conditions – e.g. fits with conversion parts and temperatures – it is possible also to use the other radial-play categories C2 (<CN) or C3 and C4 (>CN).

Axial play is derived from the radial play and raceway and rolling-element geometries, and is always given as a reference parameter.

When bearings are installed the term used is bearing clearance. During operation the bearing clearance is determined by the original bearing play, fits and material of the conversion parts such as shaft and housing, and the temperature difference between the bearing races. As a rule a perfectly running bearing should have a very small bearing clearance.

Materials

Bearing races and rolling elements are primarily composed of chrome-alloy special steel 100Cr6 (DIN EN ISO 683-17, [8]) or 52100 (ASTM A295, [9]) with a high degree of purity and hardness in the range of 58 to 65 HRC. For special applications the races and rolling elements can be made of other materials, e.g. ceramic.

Rolling-bearing cages are made of metal or plastic. The metallic cage in small rolling bearings is primarily composed of sheet steel. Polyamide 66 (PA66) is used for the majority of plastic cages. This material, especially when reinforced with fiberglass, is characterized by its favorable combination of strength and elasticity.

Other thermoplastics and duroplastics are also used as cage materials for special applications subject to extremely high thermal loads.

Selection of rolling bearings

It is necessary to take into account many external factors in order to choose the right bearing from the wealth of options.

Selection criteria*Load*

The size and direction of the load acting on the rolling bearing normally determine the type and size of the bearing. Deep-groove ball bearings are usually used in the case of low to medium loads. Rolling bearings have one advantage in the case of high loads and limited installation space. With the exception of only purely radially loaded needle bearings, cylindrical rolling bearings and axial bearings, rolling bearings can simultaneously accommodate radial and axial loads (combined loads).

Deep-groove ball bearings transfer axial loads in both directions, while angular-contact ball bearings and tapered rolling bearings can only be axially loaded in one direction.

Cylindrical rolling bearings and self-aligning bearings are particularly suitable for radial loads, but less so for axial loads.

Speed limit

Ball bearings with point contact between rolling elements and raceways have a higher speed limit than rolling bearings of the same size. The permissible speed of a rolling bearing is dependent above all on the type of construction, the size, and also the lubricating process. A bearing lubricated with oil usually has a higher speed limit than a bearing lubricated with grease.

Mounting

A distinction is made between locking bearings (completely assembled) and non-locking bearings (can be disassembled). Non-locking bearings include tapered rolling bearings, angular-contact ball bearings, cylindrical rolling bearings, and needle bearings. These bearings are for the most part easier to assemble and disassemble than locking bearings such as deep-groove ball bearings and self-aligning bearings. During mounting on the shaft and in the housing, tapered rolling bearings and angular-contact ball

bearings must be adjusted to the required bearing clearance and preload, which always calls for great care.

Further selection criteria

In addition to the previously mentioned main criteria, it is also necessary to take into account – when choosing a rolling bearing – angular adjustability to compensate for misalignment between the bearing points, running smoothness, friction, and costs.

Layout of bearing arrangement

As a rule, two bearings arranged at a specific distance from each other are required to guide and support a rotating machine part. There are two important variations of bearing arrangement: The locating/floating bearing arrangement and the preloaded bearing arrangement.

Locating/floating bearing arrangement

Two radial bearings are seated on the shaft and in the housing (Figure 2). The distance between the two bearing points is determined by manufactured conversion parts within the framework of the tolerance. In addition, the shaft, when

heated to different temperatures or when made from different materials, does not expand in the same way as the housing. These differences must be compensated for in the bearing points. For this reason, one bearing is to be axially secured as a locating bearing on the shaft and in the housing and one bearing is to be movable in the axial direction as a floating bearing. Typical applications of the locating/floating bearing arrangement can be found in alternators and steering motors.

Single-row deep-groove ball bearings are frequently used as locating bearings. Cylindrical rolling bearings, needle bearings, and deep-groove ball bearings are usually used as floating bearings.

Double-row angular-contact ball bearings and tapered rolling bearings are also used, e.g. as wheel bearings, in the case of high radial and axial loads.

Preloaded bearing arrangement

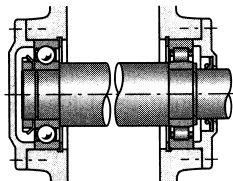
A preloaded bearing arrangement is predominantly formed from two angular-contact ball bearings or tapered rolling bearings arranged in mirror-image fashion (Figure 3). During assembly, a bearing race is displaced on its seat surface until the bearing arrangement has attained the desired or required play or preload. Because of the possibility of play regulation, a preloaded bearing arrangement is particularly suitable for applications with close guidance, e.g. bearings in transmissions.

Tolerances and fit of bearing points

Rolling bearings essentially have negative tolerances for hole diameter, outside diameter and width, i.e. the nominal size is always the maximum limit of size.

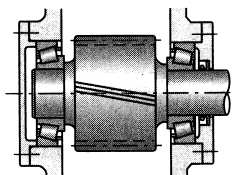
Mounting the races on the bearing points (shaft and housing bore) is important for the installation of rolling bearings. The rolling bearings must not twist, above all tangentially, under load on the counterparts. The safest and easiest way of ensuring mounting is to make the correct choice of fit and tolerances so that the load capacity of the bearing can be fully utilized. Depending on the tolerance band of the bearing seat, the fit is referred to as clearance fit, transition fit, or interference fit (Figure 4).

Figure 2: Locating/floating bearing arrangement



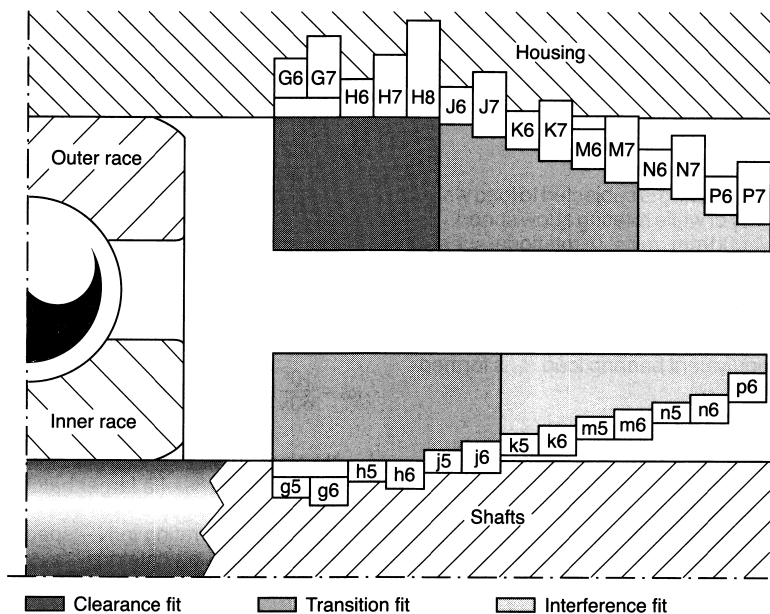
UAM0118Y

Figure 3: Preloaded bearing arrangement



UAM0119Y

Figure 4: Fits of bearing seat



UAM0120E

The load conditions of the bearing races are of great importance to the choice of fit. A distinction is made between two loads, based on the load direction and the rotation of the bearing races.

- Rotating load: The race rotates relative to the load direction, and must be firmly seated.
- Concentrated load: The race is stationary relative to the load direction, and can have a close clearance fit to interference fit with its counterpart.

Because of the small thicknesses of the bearing races, the form variations of the seats are transferred to the raceways. The counterparts should therefore demonstrate as much form quality as possible, such as concentricity, cylindricity, and runout.

Calculation of load capacity

It is necessary to distinguish between static and dynamic load capacity when calculating the load capacity of a rolling bearing. The basic principles for calculating the static and dynamic load capacities are set out in ISO 76 [10] and ISO 281 [11].

Static load capacity

If a rolling bearing is subjected to load while stationary or while rotating at low speed, i.e. $n d_m \leq 4,000 \text{ mm} \cdot \text{min}^{-1}$ (n rotational speed, d_m mean value of bore and outside diameters), it is considered under static load.

If a bearing is subjected to load in both the radial and the axial directions, the static equivalent bearing load P_0 is formed from this:

$$P_0 = X_0 F_r + Y_0 F_a \text{ in N}$$

where

X_0 Radial factor; $X_0 = 0.6$ for single-row deep-groove ball bearing.

Y_0 Axial factor; $Y_0 = 0.5$ for single-row deep-groove ball bearing.

F_r Radial load in N.

F_a Axial load in N.

At $P_0 < F_r$, $P_0 = F_r$ must be reckoned on. For other bearing types, refer to the specifications in the catalogs of rolling-bearing manufacturers for the factors X_0 and Y_0 .

As the measure of the static load capacity, a ratio

$$f_s = \frac{C_0}{P_0}$$

is formed, where C_0 is termed the static load rating. C_0 is the load at which permanent total plastic deformation of rolling elements and raceways at the contact point subject to greatest load amounts to 0.0001 of the rolling-element diameter. Catalogs specify C_0 for all rolling bearings. The calculation of the static load rating C_0 is shown in the standard ISO 76.

In the case of normal requirements, the characteristic figure $f_s = 1$ can be approved. The requirement of lower deformation (< 0.0001) corresponds to a greater characteristic figure $f_s > 1$.

Dynamic load capacity

Calculating the dynamic load capacity is based on the method of the standard ISO 281. This refers to the service life of a rotating rolling bearing under load, where the running surfaces can be subject to material fatigue. The important characteristic figure of load capacity is the dynamic load rating C . This specifies the load of the rolling bearing for a nominal bearing service life of one million revolutions.

For the purpose of calculating the nominal service life, the following applies in accordance with ISO 281:

$$L_{10} = 10^6 \left(\frac{C}{P} \right)^p \text{ in revolutions,}$$

$$L_{10h} = \frac{10^6}{60n} \left(\frac{C}{P} \right)^p \text{ in hours.}$$

L_{10} Nominal service life achieved or exceeded by 90 % of a larger batch of identical bearings.

C Dynamic load rating in N – specified in rolling-bearing catalogs.

P Dynamic equivalent bearing load in N.

p Exponent, $p = 3$ for ball bearings, $p = 10/3$ for rolling bearings.

n Rotational speed in min^{-1} .

The dynamic equivalent load P is defined as the imaginary load constant in size and direction which influences the service life, such as the actually acting radial load and axial load. It is calculated from:

$$P = X F_r + Y F_a \text{ in N}$$

where

F_r Radial load,

F_a Axial load,

X Radial factor,

Y Axial factor.

The radial factor X and axial factor Y are dependent on bearing type, size, bearing clearance, and load ratio, and are specified in ISO 281 and in rolling-bearing catalogs.

Modified service life

In addition to the nominal service life, an extended modified service life L_{na} has been introduced in the standard ISO 281, for which operating conditions can also be included in the calculation:

$$L_{na} = a_1 a_2 a_3 L_{10},$$

where

- a_1 Coefficient for survival probability,
e.g.:
90 %: $a_1 = 1$;
95 %: $a_1 = 0.62$.
- a_2 Coefficient for special bearing design
such as internal construction and materials.
- a_3 Coefficient for operating conditions,
such as bearing lubrication and operating temperature.

The coefficients a_2 and a_3 are not mutually independent, and are frequently combined as coefficient a_{23} :

$$L_{na} = a_1 a_{23} L_{10}.$$

Numerous systematic examinations and experiences from practice make it possible to quantify the influences of bearing materials and operating conditions on the achievable service life of rolling bearings. Diagrams and computing programs are provided by rolling-bearing manufacturers to calculate the coefficients a_2 , a_3 and a_{23} .

References

- [1] DIN 623: Rolling bearings – Fundamental principles:
Part 1: Designation, marking.
Part 2: Graphical representation of rolling bearings.
- [2] ISO 492: Rolling bearings – Radial bearings – Dimensional and geometrical tolerances.
- [3] DIN 620-1: Rolling bearings; gauging methods for dimensional and running tolerances.
- [4] DIN 620-2: Rolling bearings; tolerances for rolling bearings; tolerances for radial bearings.
- [5] DIN 620-3: Rolling bearings; tolerances for thrust bearings.
- [6] DIN 620-4: Rolling bearings – Rolling bearing tolerances – Part 4: Radial internal clearance.
- [7] DIN 620-6: Rolling bearings – Rolling bearing tolerances – Part 6: Chamfer dimension limits.
- [8] DIN EN ISO 683-17: Heat-treated steels, alloyed steels and free-cutting steels – Part 17: Ball and roller bearing steels.
- [9] ASTM A295: High-Carbon Anti-Friction Bearing Steel.
- [10] ISO 76: Rolling bearings – Static load ratings.
- [11] ISO 281: Rolling bearings – Dynamic load ratings and rating life.

Seals

Seal technology

Function

The function of a seal is to separate two different media from each other. For economic reasons it is necessary to seal existing sealing points with simple geometries. The O-ring assumes the simplest of these geometries.

Classification

A distinction is made between static and dynamic seals (Figure 1). In the case of a static seal there is no relative movement between seal and machine parts; in the case of a dynamic seal there is relative movement between seal and machine parts. This can exist both between seal and shaft and between seal and housing.

100 % tightness can be achieved with static seals, but this for the most part cannot be achieved with dynamic seals. A "leakage rate", which may not be exceeded, is defined for dynamic seals. When a seal complies with this rate, it is said to be subject to technical tightness.

Materials for seals

In the main elastomers but also other materials are used for seals in automotive applications. Elastomers are used in a diverse range of applications – as tires, in door seals, as axle boots, in fuel hoses, or as assorted seals in the engine compartment.

Requirements

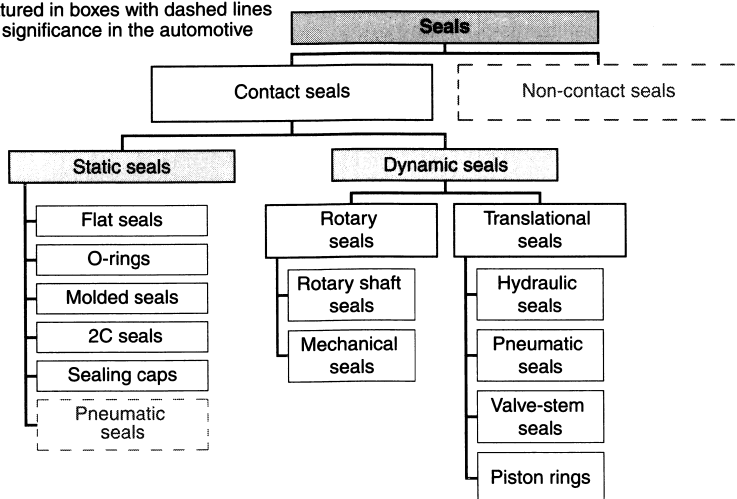
Elastomers, like all organic-chemicals, are not subject to unlimited use. External influences, e.g., different media, oxygen or ozone as well as pressure and temperature, alter the material properties and hence their behavior. Elastomers can swell, shrink, harden, crack, or even break.

The complete sealing system must always be taken into consideration in the design of a new seal. This comprises the following seven points:

- seal geometry,
- system pressure (averaged pressure and short-time peak pressures),

Figure 1: Types of seals

Seals featured in boxes with dashed lines are of no significance in the automotive sector.



- system temperature (averaged temperature and short-time peak temperatures),
- gap to be sealed,
- roughness of the mating face,
- medium to be sealed,
- relative speed between seal and machine part.

Applications

The matching sealing system is designed in line with the prevailing requirements. An overview of the existing solutions and where they are used is provided in Figure 2.

Figure 2: Seal applications in a motor vehicle

1 Air-conditioning system

- O-rings
- PTFE seals
- Rotary seals

2 Heat management

- Special rotary seals
- Multi-component seals
- Sealing rings
- O-rings
- PTFE components

3 Cylinder head

- Metal gaskets
- Multi-layer seals

4 Fuel-injection system

- O-rings
- Customer-specific sealing rings
- PTFE molded rings

5 Exhaust-gas recirculation

- EGR seal

6 Battery

- Elastomer pressure-limiting valve

7 Safety-relevant components (airbag)

- Customized two-component parts

8 Electrics

- Customized housing cover made of thermoplastic material or in combination with silicone

9 Electronic control units

- Specially developed multi-layer seals

10 Engine, transmission, steering, among others

- sealing caps

11 Brake system

- O-rings
- Customer-specific molded parts

12 Drivetrain and transmission

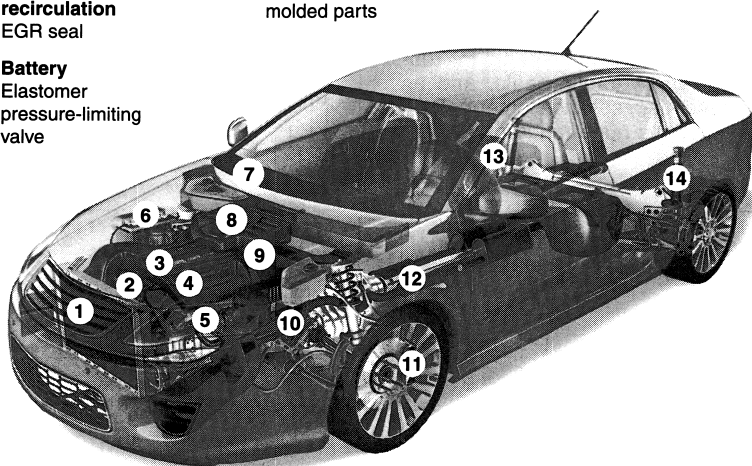
- Rotary seals
- Guide rings
- Back-up rings
- O-rings
- Sealing washers
- Sealing caps

13 Windows and doors

- Extruded elastomer sections

14 Chassis systems

- Sealing-edge ring
- O-rings
- Buffer ring
- Back-up rings
- Vane seal



O-ring

With the O-ring, design engineers are provided with an efficient and economic sealing element for a huge variety of different applications. It is used primarily as a static sealing element, but can also be used in dynamic applications. Cost-effective manufacturing processes and easy handling make the O-ring the most commonly used seal.

A large selection of elastomer materials for standard and special applications enables virtually all liquid and gaseous media to be sealed.

Description

O-rings are continuously vulcanized in molds. They are characterized by their ring shape with a circular cross-section. The O-ring is defined in its dimensions by the inside diameter d_1 and the cross-section d_2 (Figure 1).

Thanks to its simple geometry the O-ring has many benefits:

- symmetrical cross-section,
- simple, compact design,
- self- and double-acting,
- simple calculation and definition of the groove,
- continuous groove design,
- large choice of materials,
- wide range of applications.

Applications

O-rings are used as primary sealing elements, as tensioning elements for rubber-energized hydraulic seals, and as wipers (secondary sealing element). They therefore cover a huge range of applications.

Whether it be as an individual seal for a repair or as a quality-assured sealing element in automobile manufacturing or mechanical engineering – there is today no field in industry in which the O-ring is not used. The O-ring is used mainly in static seals

- as radial-static sealing, e.g., in bushings, covers, pipes, and cylinders (Figure 2),
- as axial-static sealing, e.g., in flanges, plates, and plugs (Figure 3).

Dynamic use is only possible at low load levels. It is limited by the speed and the pressure to be sealed

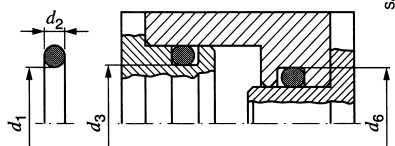
- to sealing e.g. back- and forward-moving pistons, rods, and plungers,
- to sealing e.g. slowly slewing, rotating or screw-like movements on shafts, spindles, and rotary transmission lead-throughs.

Operating concept

O-rings are self- and double-acting sealing elements. The pressure forces caused by installation in the radial or axial direction produce the initial tightness. They are

Figure 2: Radial installation of an O-ring

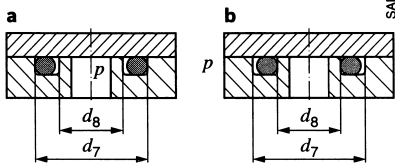
d_1 Inside diameter,
 d_2 Cross-section,
 d_3 Groove-base diameter,
 d_6 Installation-space outside diameter.



SAM0158Y

Figure 3: Axial installation of an O-ring

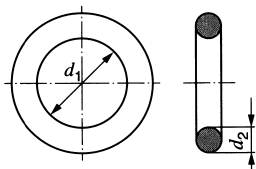
a) Pressure p from inside,
 b) Pressure p from outside.
 d_7 Groove outside diameter,
 d_8 Groove inside diameter.



SAM0159-1Y

Figure 1: O-ring dimensions

d_1 Inside diameter,
 d_2 Cross-section.

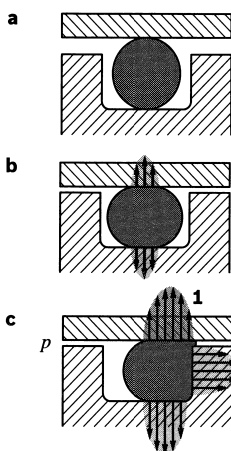


SAM0157Y

superimposed by the system pressure. This creates a total sealing pressure, which increases as the system pressure increases (Figure 4). The O-ring behaves under pressure similarly to a liquid with high surface tension. In this way, the pressure is uniformly transmitted to all sides.

Figure 4: Pressure forces on the O-ring

- a) Without pre-pressure,
 b) Pressure forces with pre-pressure,
 c) Pressure forces with pre-pressure,
 with system pressure.
 1 Pressure path in the O-ring.
 p System pressure.



SAM0160-1Y

Materials

Equipment manufacturers and operators expect sealing systems to function without leaks and to have long service lives. In the interests therefore of finding the ideal seal solution in an individual application, the choice of materials is, as well as the correct design, of crucial importance. The specific material properties will be discussed elsewhere (see Elastomers), since this apply equally to all seals.

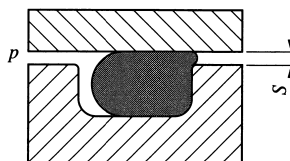
Sealing gap

Excessively large sealing gaps run the risk of gap extrusion (Figure 5), which can cause the O-ring to be destroyed.

The permissible radial sealing gap S between the parts to be sealed is dependent on the system pressure, the cross-section, the media temperature, and the Shore hardness of the O-ring. Table 1 sets out recommendations for the permissible gap dimension S as a

Figure 5: Sealing gap

- p System pressure,
 S Sealing gap.

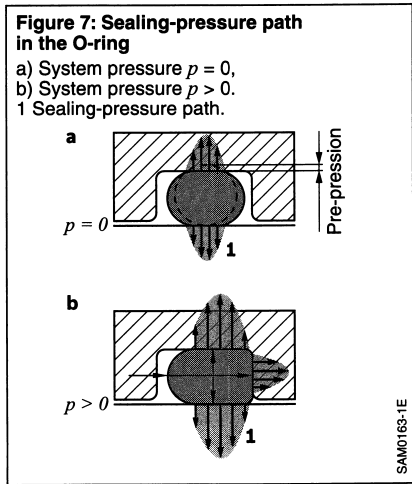
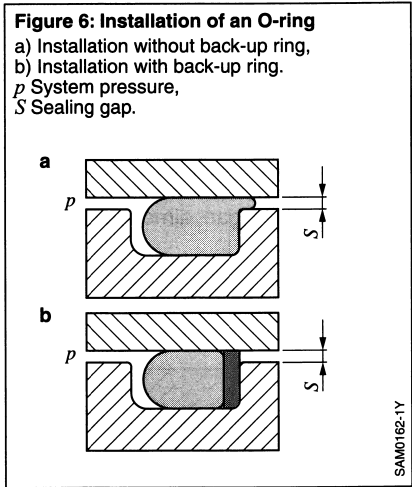


SAM0161-1Y

Table 1: Gap dimensions

O-ring cross-section d_2 [mm]	up to 2	2...3	3...5	5...7	>7
O-rings with hardness 70 Shore A					
Pressure p [MPa]	Gap S [mm]				
≤ 3.50	0.08	0.09	0.10	0.13	0.15
≤ 7.00	0.05	0.07	0.08	0.09	0.10
≤ 10.50	0.03	0.04	0.05	0.07	0.08
O-rings with hardness 90 Shore A					
Pressure p [MPa]	Gap S [mm]				
≤ 3.50	0.13	0.15	0.20	0.23	0.25
≤ 7.00	0.10	0.13	0.15	0.18	0.20
≤ 10.50	0.07	0.09	0.10	0.13	0.15
≤ 14.00	0.05	0.07	0.08	0.09	0.10
≤ 17.50	0.04	0.05	0.07	0.08	0.09
≤ 21.00	0.03	0.04	0.05	0.07	0.08
≤ 35.00	0.02	0.03	0.03	0.04	0.04

function of the O-ring cross-section and the Shore hardness. The table applies to elastomer materials, except polyurethane and FEP- or PFA-sheathed O-rings. In the case of pressures greater than 5 MPa for inside diameters larger than 50 mm and pressures greater than 10 MPa for inside diameters smaller than 50 mm, it is necessary to provide back-up rings (Figure 6).



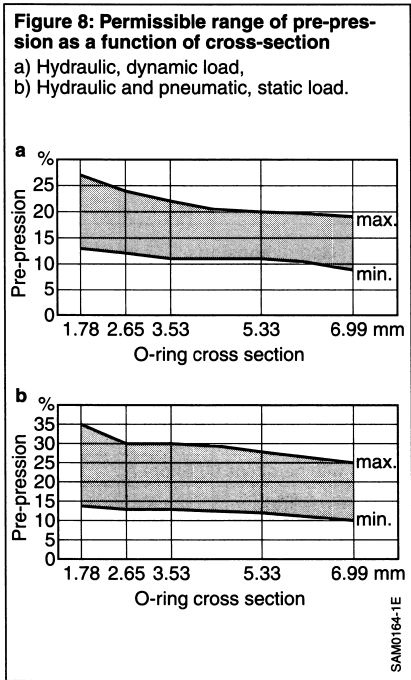
Groove filling
To avoid damaging effects on the sealing function, it is important to take into consideration the groove filling for an installed O-ring. It should in the installed state where possible not exceed 85 % so as to accommodate any thermal expansion of the O-ring, volume swelling due to media contract, and influences of tolerances.

Pre-pressure
The pre-pressure (Figure 7) serves among other things

- to achieve the initial tightness,
- to bridge production-related tolerances,
- to ensure defined frictional forces,
- to compensate the compression set,
- to provide compensation for wear.

Depending on the application, the following values, referred to the cross-section (d_2), are recommended for the pre-pressure: for dynamic use 6 to 20 %, for static use 15 to 30 %.

For the design of grooves the guide values for pre-pressure can be taken from the



diagrams in Figure 8 and Figure 9. These taken into account in accordance with ISO 3601-2 [1] the dependence on loads and the cross-section.

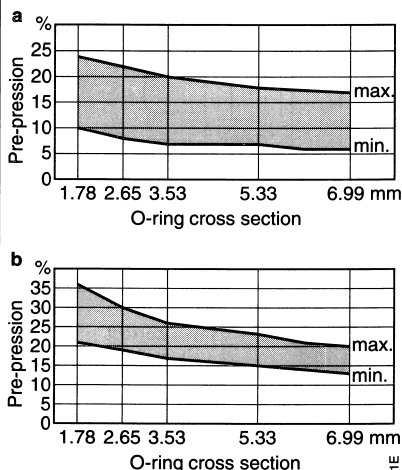
Surfaces

Elastomers adapt to irregular surfaces under pressure. For gas- or liquid-tight connections, however, minimum demands must be made of the surface quality of the surfaces to be sealed. Basically, for instance, scores, scratches, shrink holes, and concentric or spiral machining marks are not permitted.

Dynamic mating faces are subject in terms of surface quality to more stringent demands than static seals. There are as yet no standardized definitions for describing mating faces. The specification of the R_a value (roughness average) is in practice not sufficient to assess the surface quality. Manufacturer recommendations therefore contain different terms and definitions among others in accordance with DIN 4768 [2] and DIN EN ISO 4287 [3].

Figure 9: Permissible range of pre-pressure as a function of cross-section

- a) Pneumatic, dynamic load,
b) Axial-static load.



General technical data

O-rings can be used in a wide range of applications. Temperature, pressure and media determine the choice of suitable materials. To be able to assess the suitability of the O-ring as a sealing element for a given application, it is necessary to consider the interaction of all the operating parameters. Further information can also be obtained in the calculation program for O-rings of Trelleborg Sealing Solutions [4].

Usage criteria

Operating pressure

Static usage

In static usage the following values apply to the operating pressure:

- operating pressure < 5 MPa for inside diameters > 50 mm without back-up ring.
- operating pressure < 10 MPa for inside diameters < 50 mm without back-up ring (depending on material, cross-section and gap dimension).
- operating pressure < 40 MPa with back-up ring.
- operating pressure < 250 MPa with special back-up ring.

The permissible gap dimensions must be observed here.

Dynamic usage

In dynamic usage the following values apply to the operating pressure:

- operating pressure < 5 MPa in back-and forward-moving usage, without back-up ring.
- higher pressures with back-up ring.

Speed

In back- and forward-moving usage speeds of up to 0.5 m/s are permitted, in the case of rotating seals also up to 0.5 m/s. This applies in each case depending on material and application.

Temperature

Depending on material and media resistance, usage in the temperature range of –60 to +325 °C is permissible.

The short-time peak and continued-use temperatures and the operating time must be taken into account when assessing the

usage criteria. In the case of rotating usage the temperature increases caused by frictional heat must be observed.

Media

Virtually all liquids, gases and chemicals can be sealed with a large variety of materials with different properties. Details on the choice of suitable materials can be found elsewhere (see Elastomers). Further information can be found on the Internet, e.g. in the Chemical Compatibility database [5].

Notes on installation

Before starting installation, it is necessary to check the following points:

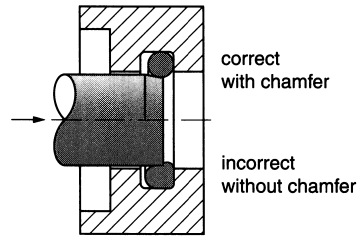
- Lead-in chamfers fashioned as per drawing?
- Internal bores deburred and rounded?
- Machining residues (e.g., swarf, dirt, foreign particles) removed?
- Thread crests covered?
- Seals and components greased or oiled? Media compatibility with elastomer must be ensured. It is recommended to use the medium to be sealed for lubrication. No lubricants with solid additives (e.g., molybdenum disulfide, zinc sulfide) may be used.

An installation-oriented design can eliminate potential causes of seal failure (examples in Figures 10, 11 and 12). Since O-rings are always installed with interference, lead-in chamfers and edge roundings must be provided.

Installation by hand

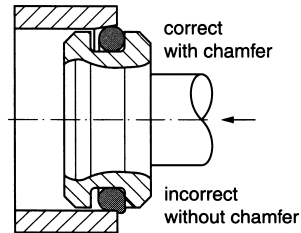
- Do not use any sharp objects.
- Look out for twisting, use aids and appliances for correct positioning.
- Use fitting aids wherever possible.
- Do not overstretch O-rings.
- Do not stretch O-rings made of extruded round cord over the joint.

Figure 10: Rod installation with O-ring



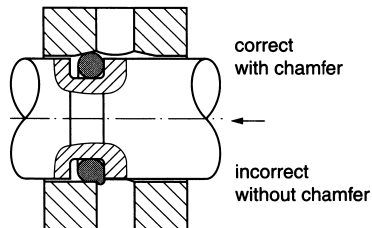
SAM0166E

Figure 11: Piston installation with O-ring



SAM0167E

Figure 12: O-ring installation over cross-bores



SAM0168E

Installation over threads, shafts, and similar

If the O-ring has to be guided during installation over threads, shafts, keyways or similar, a mounting sleeve will be required. This should not have any sharp edges or burrs and can be made of soft metal or plastic.

Lead-in chamfers

The minimum lengths of the lead-in chamfers are specified in Table 2 as a function of the cross-section d_2 (see also Figure 13 and Figure 14). The surface roughness of the lead-in chamfer is specified as:

$$R_z \leq 6.3 \mu\text{m}; R_a \leq 0.8 \mu\text{m}.$$

Installation types and information on installation-space design

Installation types

O-rings can be used in a variety of applications in components. The subsequent fitting situation must already be taken into account in the design. Sharp edges and bores should be circumvented so as to avoid damage during fitting. In the case of long sliding movements, the seal seat must if possible be stepped or the O-rings arranged in such a way that they only cover short mounting distances. Otherwise there is the risk of twisting.

Radial installation (static and dynamic)

The O-ring size in the case of rod seals (internally sealing) must be chosen in such a way that the O-ring outside diameter ($d_1 + 2 d_2$) is at least equal to or greater than the installation-space outside diameter d_6 (Figure 2).

The O-ring size in the case of piston seals (externally sealing) must be chosen

in such a way that the inside diameter d_1 is equal to or less than the groove-base diameter d_3 (Figure 2).

Axial installation (static)

In the case of axial-static installation, the pressure direction must be taken into account in the choice of the O-ring size (Figure 3). For internal pressure the O-ring outside diameter should be chosen as being equal to or greater than the groove outside diameter d_7 . For external pressure the O-ring inside diameter should be chosen as being less than the groove inside diameter d_8 .

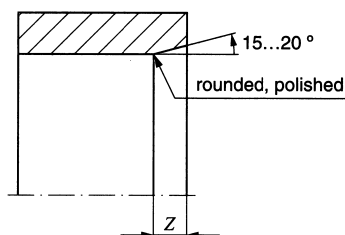
Elongation and upsetting

Radial installation of piston seal

When the O-ring is used as a piston seal (externally sealing), the nominal inside diameter d_1 of the O-ring (see Figure 2) should be elongated for dynamic applica-

Figure 13: Lead-in chamfer for bores and pipes

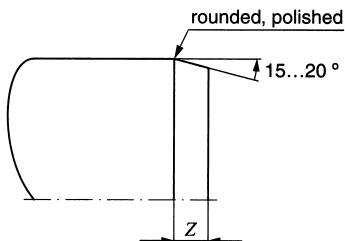
Z Lead-in chamfer.



SAM0169-1E

Figure 14: Lead-in chamfer for shafts and rods

Z Lead-in chamfer.



SAM0170-1E

Table 2: Lead-in chamfers

Lead-in chamfers Length Z min. [mm]		O-ring cross-section d_2 [mm]
		Standard dimensions converted from inches (metric standard dimensions)
15 °	20 °	
2.5	1.5	up to 1.78 (1.80)
3.0	2.0	up to 2.62 (2.65)
3.5	2.5	up to 3.53 (3.55)
4.5	3.5	up to 5.33 (5.30)
5.0	4.0	up to 7.00
6.0	4.5	over 7.00

tions between 2 % and 5 % and for static applications between 2 % and 8 %. In the case of O-rings with an inside diameter $d_1 < 20$ mm it is not always possible to maintain this, which can result in a larger elongation range. To minimize the elongation range and the maximum stretching, it is necessary to minimize the diameter in the groove base d_3 (see Figure 2) and to impose less stringent demands on minimum stretching.

In dynamic applications it is important not to exceed a maximum elongation of 5 % so as to avoid detrimental effects on the sealing function. Basically, exceeding these recommended values will lead to a greater reduction of the O-ring cross-section, which as a result may impact on the service life of the O-ring.

Radial installation of rod seal

If the O-ring is used as a rod seal (internally sealing), the outside diameter of the O-ring ($d_1 + 2d_2$) should be at least equal to or greater than the outside diameter of the installation space (groove base) d_6 (see Figure 2) so as to achieve upsetting of the O-ring outside diameter. The outside diameter of the O-ring should for O-rings with a diameter $d_1 > 250$ mm not exceed 3 % of the groove outside diameter and 5 % for O-rings with a diameter $d_1 < 250$ mm. For O-rings with a diameter $d_1 < 20$ mm, this is not always possible on account of the tolerance position, which may result in greater upsetting of the outside diameter. Basically, exceeding these recommended values will lead to a marked increase in the O-ring cross-section, which as a result may impact on the service life of the O-ring.

Axial-static installation

If the O-ring is used as an axial-static seal, the pressure direction should be taken into account in the choice of the O-ring size (Figure 3). For O-ring pressurization the O-ring size should be chosen in such a way that prior to pressurization the O-ring fits closely on the groove flank on the non-pressure side. For internal pressure the O-ring should be chosen in such a way that the outside diameter ($d_1 + 2d_2$) of the O-ring is equal to or slightly greater (max. approximately 1 to 2 %) than the groove outside diameter d_7 . For external pressure the O-ring should be approximately 1 to 3 % smaller than the groove inside diameter d_8 .

Back-up rings

Back-up rings have no sealing functions, but, as the name suggests, are protective and supporting elements made of extrusion-resistant materials with a predominantly rectangular cross-section. They are installed together with an elastomer seal, as a rule with an O-ring, into a groove for static application. The close fit between back-up ring and bore or rod prevents the O-ring, which is under pressure, from extruding into the sealing gap.

Advantages

- Use of O-rings in high-pressure applications,
- use of O-rings of low hardness,
- compensation of large radial gap dimensions,
- externally and internally sealing application possible,
- used for static and back- and forward-moving or slowly rotating movements,
- compensation of gap enlargement by thermal expansion,
- static and dynamic applications.

Molded seal

Application

Downsizing and weight reduction in vehicles are imposing increasingly more stringent demands on the sealing systems. Engine components and add-on parts are increasingly being manufactured from plastics. Owing to the lower rigidity levels, plastic components are subject to greater deformation than metal components (e.g., axial and radial tolerances, tolerances due to high pressures).

Diecast metal components are as a rule not remachined, entailing greater component tolerances accordingly. These demanding requirements can be satisfied very well with elastomer sealing systems (Figure 1).

Sealing system

Influencing variables

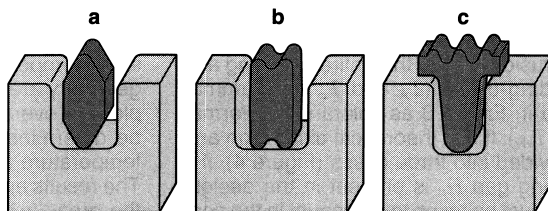
Figure 2 shows the significant influencing variables for a sealing system. It depicts a molded seal for sealing between housing and cover. The system's requirements profile can be divided into the following stress types: mechanical (pressure, vibrations), chemical or physical (fluids or media), thermal, and electrical.

Design features

When it comes to the design features, the installation space (Figure 2, A1) is considered with its tolerances, flatness, the surface texture, the sealing gap, and the radius and flank angles. The installation space consists of the seals partners groove and counter-sealing face. The profile of the molded seal (Figure 2, A2) is matched under the criteria of compression and the clearance examination to the installation space.

Figure 1: Design examples

- a) I-profile,
- b) Double-I-profile,
- c) T-profile.



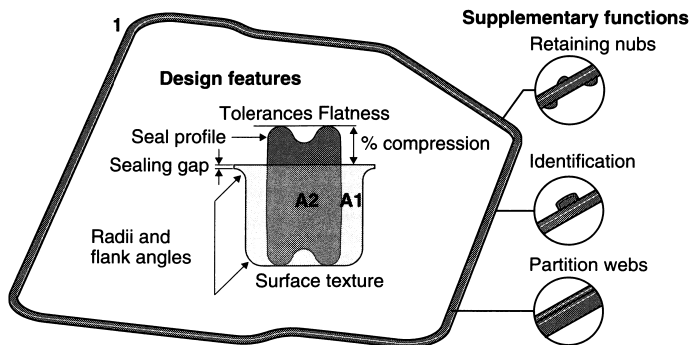
SAM0206Y

Figure 2: Significant influencing variables for a sealing system

1 Molded seal (sealing between housing and cover).

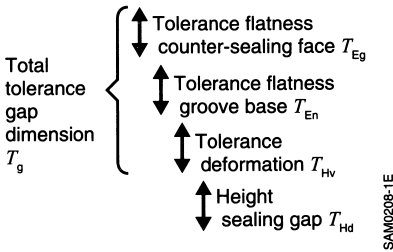
A1 Installation space,

A2 Profile of molded seal.



SAM0207-1E

Figure 3: Relationship between gap dimension and sealing gap

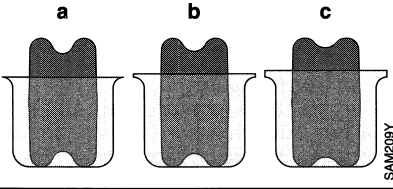


Additional functions can also be used on the molded part. Consequently, retaining nubs are used to provide transport retention for example during fitting, an identification for traceability, and partition webs for locating a leak.

The overview in Figure 3 shows the relationship between gap dimension and sealing gap. The gap dimension T_g is the sum of the component tolerances consisting of flatness of groove base T_{En} and counter-sealing face T_{Eg} , the bushing protrusion T_B and the entire widening and bending-up T_A (T_A and T_B are summarized in Figure 3 as Tolerance Deformation T_{Hv}). It is a theoretical dimension and is divided into three cases (Figure 4). If a sealing gap H_d is present in the design, this must be taken into account in the configuration of the compression.

Figure 4: Overview: minimum, nominal and maximum clearance

- a) Maximum sealing, minimum installation, minimum clearance.
- b) Nominal sealing, nominal installation, nominal clearance.
- c) Minimum sealing, maximum installation, maximum clearance.



Finite-element method

The finite-element method is a numeric analysis method for obtaining estimates for complex technical problems (see Finite-element method). All FEM calculations are based on idealized input data. A calculation is expedient only if the input variables for the material model, the system's requirements profile (e.g., pressure, temperature, installation) and the geometry are known precisely. For example, the overall elongation of a seal can be determined after installation at room temperature and a defined overpressure. The results are shown in graphic form with the areas with differing elongation being colored.

Flat seals

Application

Flat seals (gaskets) and their limits of use are very much dependent on the material used. Here, the applications vary in vehicle manufacturing from hot fluids, through strong chemical stress, to high pressures. Depending on the requirement, the flat seal can be sheathed in metal or even made up entirely of metal.

The different types of flat seal are as follows:

- Less hard-wearing: Paper, compressed fiber, elastomer, hard- and elastomer-bonded cork.
- Medium hard-wearing: Bonded glass-fiber, aramide-fiber, carbon-fiber and mineral-fiber seals (soft-material seals), elastomer-metal composite.
- Strongly hard-wearing: Expanded graphite, metal.
- Corrosion-resistant: Solid, expanded, filled PTFE; corrosion-resistant metal, elastomer-metal composite.

Variants

As the above list shows, there is a large variety of flat seals. Only the most important of this will be discussed in the following.

Cylinder-head gaskets

Cylinder-head gaskets (Figure 1) are fundamental static seals in the engine compartment. The complex flat seal as a rule consists of one or more coated metal layers with the possibility of having elastomer

areas vulcanized on. The cylinder-head gasket not only seals against combustion gas, oil and coolant, but also acts as a power-transmission element between the cylinder head and the crankcase.

Soft-material seals

Soft-material seals made of elastomer-fiber composite materials are, in terms of their basic principle, some of the oldest seals. They depend first and foremost on the quality of their fibers and are very pressure- and temperature-resistant. They are used for example as intake-manifold seals and in exhaust-gas recirculation valves. Natural rubber, NBR, SBR, CR, HNBR and FKM are used (see Elastomers) as binders in elastomer-bonded soft-material seals. For higher stresses simply flat seals with expanded graphite or pure metal gaskets can be used.

Metal-bead gaskets

In the case of narrow sealing webs or low surface pressure, a good solution is to use metal-bead gaskets (used for example as cover gaskets in pressure applications). The surface pressure is reduced on account of the beads to a linear pressure. This allows a higher surface pressure with consistent screw force. An elastomer coating applied to the carrier plate also facilitates the micro-sealing effect.

Elastomer-metal flat seal

The flat seal made of an elastomer-metal composite is particularly suitable for preventing micro-leaks. This inherently sta-

Figure 1: Example of a cylinder-head gasket for a cylinder bank of a 6-cylinder engine with two exhaust and two intake valves

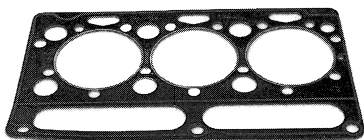
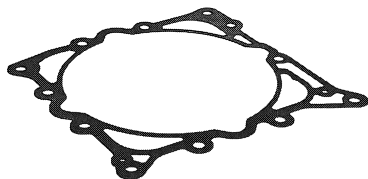


Figure 2: Flat seal in an electric-motor housing with integrated cooling



ble and at the same time adaptable seal consists of a thin metal layer with vulcanized-on elastomer layer (Figure 2). Stamped into complex geometries, the seal is ideal for use in primary frictional connection. With an elastomer-layer thickness of at least 0.2 mm, it also offers the opportunity of compensating for shrink holes and pores in the casting. The elastomer-metal flat seal is used in flange joints, pumps, automatic transmissions, and electric motors.

Sealing caps

Application

Sealing caps belong to the group of static seals. They facilitate the reliable sealing of assembly- or production-related bores at various points in the engine compartment. Design engineers can always choose between the standard solution according to DIN 443 [6] and other special types which meet the high requirements of the automotive industry.

The standard solution according to DIN 443 describes a cap made of metal which is generally pressed into the housing and provides secure sealing only by means of an additional adhesive connection. Set against these caps are sealing caps, which have an elastomer layer and do not require any additional adhesives. The elastomer layer furthermore facilitates installation without damaging the housing surface and thereby prevents micro-leak paths.

Technology

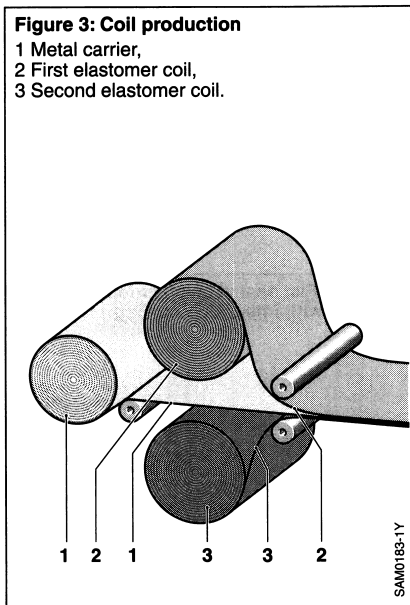
Metal sealing caps are worked in a deep-drawing process from the coil material into the finished product. The favorable elastomer-metal sealing caps can be manufactured in different ways. Possible production processes are compression molding, injection molding, lamination or vulcanization. One innovative production process involves the technology whereby multiple-layer rubber-metal connections are produced, and then stamped and deep-drawn as composite material (Figure 3).

The process of applying a binder to metal carriers and vulcanizing elastomers on metal creates secure and solid connections which can be fashioned into complex seal geometries. In this way, flat seals, seals with metal frames for automated handling, friction-optimized guide bands, and housing gaskets and sealing caps can be specifically produced.

Often, as a result of the optimized rubber-metal connection, the surfaces of the sealing faces no longer have to be machined in a complicated process. The elastomer layer improves sealing performance and, furthermore, no adhesives or sealing compounds are required in

Figure 3: Coil production

- 1 Metal carrier,
- 2 First elastomer coil,
- 3 Second elastomer coil.



production, enabling the cycle times in assembly to be reduced and the production process to be fashioned in a more process-reliable way. Users gain cost advantages through simpler handling, the possibility of automation in assembly, and reduced logistics expenditure. There are cost-reduction potentials, particularly in respect of injection-molded sealing caps.

Examples in practice

Engine blocks are produced by sand casting. At the end of the production process, the engine block features a number of bore holes which cannot be avoided from a production standpoint, but must nevertheless be sealed. Rubber-metal sealing caps offer an ideal solution to this since they are easily mounted, not bonded, and rustproof. The installation situation is pictured in Figure 4.

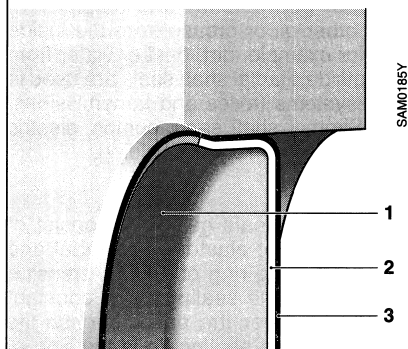
There are further possible uses in the form of static sealing of bore holes in the transmission and in steering systems. But rubber-metal sealing caps are also used in practice in the camshaft housing and crankcase and in pumps.

Variants

Rubber-metal sealing caps can be manufactured from different materials. Steel, galvanized steel and stainless steel may be used as metal materials. Possible elastomer materials are NBR, EPDM and AEM, and hence provide for a temperature range of -45 to 150 °C and different media-resistance capabilities.

Figure 4: Installation situation of sealing cap

- 1 Sealing cap,
- 2 Metal,
- 3 Rubber layer.



Radial shaft seal

Application

Radial shaft seals are ring-shaped sealing elements the purpose of which is to separate permanently and safely from each other oil or grease from the inside and, for example, dirt, dust or water from the outside. Radial shaft seals are used in drive systems (axles and transmissions, e.g., Simmer shaft seal), pumps, electric motors, and in machine-building.

Design

Radial shaft seals generally consist of a sealing lip of elastomer material and a strengthening ring of metal. A tension spring lends the sealing lip its constant pre-tension over the entire service life (Figure 1). The sealing lip is the end of the diaphragm (sealing cup). The sealing edge is the area of the sealing lip that touches the shaft and assumes the actual sealing function.

The sealing-lip geometry corresponds to today's state-of-the-art and is based on many years of application-technology experience. The sealing edge can be press-finished or manufactured at the face by mechanical cutting.

The entire radial force of the seal is created by the initial force of the elastomer sealing lip and the tensile force of the

spring. The former is obtained depending on deformation from the elasticity of the material, the sealing-lip geometry, and the overlap between the shaft and the seal.

Specifically for the automotive industry the function and application-related information regarding radial shaft seals are described in DIN 3761 [7] and in ISO 6194 [8].

Seal design

Standard types

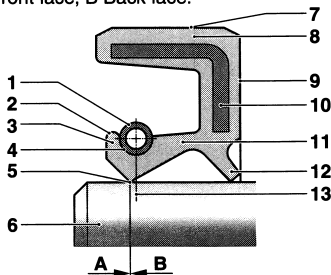
Radial shaft seals are available in standard types according to DIN 3760 [9] and DIN 3761 or ISO 6194. Figure 2 shows versions in which the strengthening ring is completely enclosed in elastomer and versions in which the strengthening ring is embedded in elastomer only in the area of the diaphragm and consequently forms the casing of the shaft seal (adhesion part to which the elastomer adheres).

Special types

Special types are used when application parameters are outside the scope of DIN 3760 and DIN 3761 or ISO 6194. This can be increased speed above 10 m/s, pressure above 0.05 MPa, grease used

Figure 1: Profile section of a radial shaft seal

- 1 Tension spring, 2 Spring-retaining lip,
 - 3 Sealing lip, 4 Spring groove,
 - 5 Sealing edge, 6 Shaft, 7 Outer casing,
 - 8 Elastomer, 9 Back casing,
 - 10 Strengthening ring,
 - 11 Diaphragm (sealing cup),
 - 12 Dust lip, 13 Spring plane.
- A Front face, B Back face.

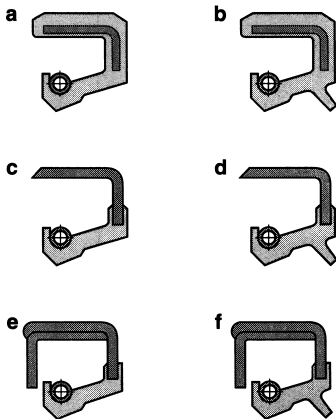


SAM0186-2Y

Figure 2: Radial shaft seals in standard types

Standard types acc. to DIN 3760 and DIN 3761 or ISO 6194.

- a) DIN 3760 Type A, b) DIN 3760 Type AS
- c) DIN 3761 Type B, d) DIN 3761 Type BS
- e) DIN 3761 Type C, f) DIN 3761 Type CS



SAM0187-1Y

instead of oil as the medium to be sealed, or increased dirt accumulation. It is recommended to use a special seal in all these situations. Figure 3 shows some special types by way of example.

Main features

The main features to be observed in any choice of seal is the outer casing, the strengthening ring, the tension spring, and the material.

Outer casing

The outer casing can be either smooth or grooved; in both cases the seal can be pressed into bores in accordance with ISO fit H8. The outside diameter of the radial shaft seal is manufactured in accordance with ISO 6194-1 (Table 1).

Metal strengthening ring

Cold-rolled sheet steel in accordance with DIN EN 10139 [10] is used for the standard version. Depending on installation conditions and environmental conditions, however, other materials such as brass and stainless steel (steel grade 1.4301 in accordance with DIN EN 10088-3 [11]) may come into play.

The primary purpose of the strengthening ring is to stiffen and reinforce the

radial shaft seal. Normally, this ring may not be subjected to axial load. Should this be necessary, a special version of the strengthening ring may also be used (e.g., metal version, partly rubberized version).

Tension spring

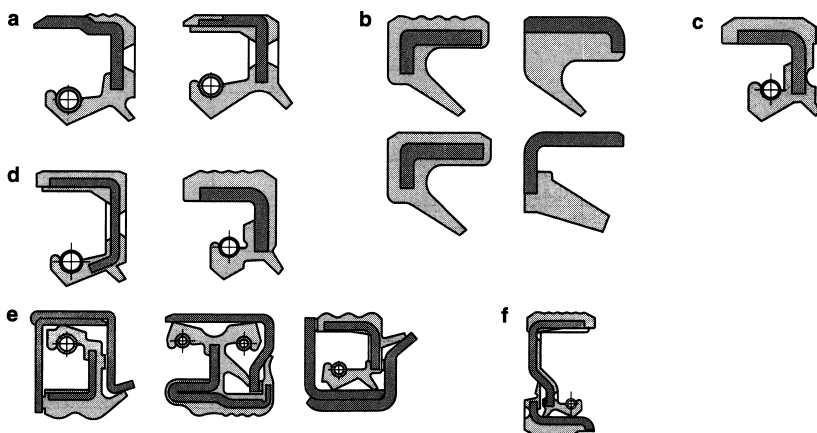
When rubber is exposed to heat, load or chemical stress, it gradually loses its original properties. This is called rubber aging. The original radial force of the sealing cup is lost. The primary purpose of the tension

Table 1: Tolerances acc. to ISO 6194-1

Outside diameter d_2 nominal [mm]	Diametrical tolerances [mm]	
	Metal housing	Rubber- coated
$d_2 < 50$	+0.20 +0.08	+0.30 +0.15
$50 < d_2 < 80$	+0.23 +0.09	+0.35 +0.20
$80 < d_2 < 120$	+0.25 +0.10	+0.35 +0.20
$120 < d_2 < 180$	+0.28 +0.12	+0.45 +0.25
$180 < d_2 < 300$	+0.35 +0.15	+0.45 +0.25
$300 < d_2 < 500$	+0.45 +0.20	+0.55 +0.30

Figure 3: Special types of radial shaft seals

- a) Types with semi-rubberized outside diameter, b) Rotary seals without tension spring, c) Type for medium to high pressures, d) Types for medium pressures, e) Cassette seals, f) System-shaft seals.



spring therefore is to maintain the radial force.

Tests have shown that the radial force must be different depending on size range and sealing-ring type. It has also emerged from these tests that it is very important to keep the deviations of radial force during the life of the seal within tight limits. The radial force has been established by comprehensive laboratory tests.

The tension spring is coiled tightly and with pre-tension. The total force of the tension spring thus comprises partly the pre-tension and partly the force obtained from the spring rate. Using a tension spring with pre-tension offers the following advantages:

- In the event of sealing-lip wear, the part of the total radial force resulting from the spring pre-tension remains unchanged.
- Through partial softening of the spring (by heat treatment) the pre-tension can be regulated in such a way that the envisaged radial force is achieved for the respective shaft diameter.

Elastomer material of the seal

The main features for ensuring reliable functioning of radial shaft seals are both the correct design and the choice of the correct material. Therefore the choice of material must be taken into consideration as one of the most important decisions with regard to media compatibility and environmental conditions. Some material properties which are directly connected to the environmental conditions are:

- good chemical resistance,
- good resistance to heat and cold,
- good ozone and weather resistance.

Functional material requirements are among others:

- high resistance to wear,
- low friction,
- low deformation by compression,
- good elasticity.

For cost reasons, good workability is desirable as an additional property. None of the materials available today can satisfy all these requirements. The choice of material therefore always constitutes a compromise between the relative importance of the respective factors.

Table 2: Material recommendations

Materials for sealing conventional media		Material designation				
		NBR	FKM	ACM	VMQ	HNBR
		Material code				
		N	V	A	S	M
		Max. permissible continuous temperature [°C]				
Mineral lubricants	Engine oils	100	170	125	150	130
	Transmission oils	80	150	125	130	110
	Hypoid gear oils	80	150	125		110
	ATF fluids	100	170	125		130
	Pressure fluids (DIN 51524, [14])	90	150	120		130
	Greases	90				100
Fire-resistant pressure fluids (VDMA 24317, [12]) (DIN 24320, [13])	Oil-water emulsion	70			60	70
	Water-oil emulsion	70			60	70
	Aqueous solution	70				70
	Anhydrous fluids		150			
Other media	Heating fuel oils	90				100
	Water	90	100			100
	Suds	90	100			100
	Air	100	200	150	200	130

NBR Nitrile butadiene rubber, FKM Fluorine rubber, ACM Acrylate rubber, VMQ Silicone rubber. HNBR Hydrogenated nitrile butadiene rubber.

Materials and their designations are

- nitrile butadiene rubber (NBR),
- acrylate rubber (ACM),
- silicone rubber (VMQ),
- fluorine rubber (FKM),
- hydrogenated nitrile butadiene rubber (HNBR).

Hydrogenated nitrile butadiene rubber (HNBR) is a further development of conventional nitrile butadiene rubber (NBR). This material offers substantially improved heat and ozone resistance and can be used instead of acrylate rubber and in certain cases also instead of fluorine rubber. A special composition is developed for each rubber type in order to satisfy the requirements imposed on seals. In addition, still more compounds are available for some extreme conditions.

Table 2 sets out material recommendations for various applications.

Design of shaft and bore

Surface texture, hardness and machining method

The shaft design is of crucial importance both to the seal and to the service life (Figure 4). Basically, the higher the circumferential speeds the greater the hardness of the shaft should be. It is established in DIN 3760 that the shaft must exhibit a hardness of at least 45 HRC. As the circumferential speed increases, the demand for hardness rises, and at 10 m/s a hardness of 60 HRC is required.

The choice of appropriate hardness is not only dependent on the circumferential speed, but also influenced by factors such as lubrication and particles that cause wear. Poor lubrication and difficult external conditions (e.g., high dust deposits on construction machinery) therefore require a higher shaft hardness. DIN 3760 and DIN 3761 lay down maximum values for the surface roughness; a surface roughness of $R_t = 1$ to $4 \mu\text{m}$ (maximum roughness) is recommended. It has, on the other hand, emerged from laboratory tests that the most favorable roughness is $R_t = 2 \mu\text{m}$ (roughness average $R_a = 0.3 \mu\text{m}$). Both rougher and finer surfaces cause higher friction, which results in higher temperature and increased wear. Consequently,

a roughness of $R_t = 2$ to $3 \mu\text{m}$ ($R_a = 0.3$ to $0.5 \mu\text{m}$) is to be recommended.

Friction and temperature measurements have also shown that grinding the shaft is the best machining method. Spiral grinding marks can however cause a pumping effect and leakage, which is why plunge-cut grinding should be chosen. Whole-number conditions of wheel speed to workpiece speed must be avoided here. Polishing the contact surface with emery cloth produces a surface texture which causes higher friction and temperature build-up than plunge-cut grinding.

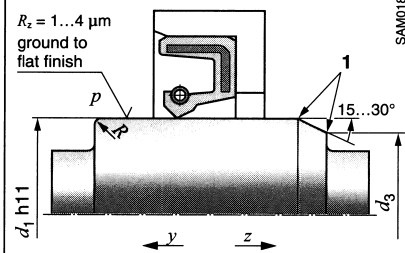
In some cases it is not possible to furnish a shaft with the hardness, surface quality and corrosion resistance required for the seal. However, this problem can be solved by fitting a separate sleeve on the shaft. In the event of wear, only the sleeve has to be replaced.

Radial run-out

Radial run-out of the shaft is to be avoided wherever possible or kept within the tightest limits. At high rotational speeds the sealing lip is at risk of not being able to follow the shaft due to its inertia. The shaft seal must be arranged in direct proximity to the bearing (e.g., roller bearing) and the bearing clearance kept as small as possible (Figure 5).

Figure 4: Shaft design

- 1 Edge rounded and polished, lead-in chamfer 15 to 30°.
- d_1 Diameter of shaft, tolerance h11,
- d_3 Inside diameter of chamfer,
- R Radius,
- R_z Average peak-to-valley height,
- p Pressure of medium,
- y, z Direction of installation.



Deviation in concentricity

A deviation in concentricity between the shaft and the accommodating bore is to be avoided wherever possible so that the sealing lip is not subjected to load on one side (Figure 6). Depending on the direction of installation *y* or *z*, it is recommended to incorporate a chamfer or a radius. The dimensions for this are shown in Figure 4 and Table 3.

Properties of the shaft surface

The values of the moving surface are established for shaft seals in DIN 3760 and DIN 3761 or ISO 6194. The surface should exhibit the following properties:

Surface roughness

$R_a = 0.2$ to $0.5\text{ }\mu\text{m}$ (arithmetic roughness average),
 $R_z = 1$ to $4\text{ }\mu\text{m}$ (average peak-to-valley height),
 $R_{\text{max}} (= R_t) = 6.3\text{ }\mu\text{m}$.

Table 3: Chamfer length for the shaft end

Diameter d_1	Diameter d_3	Radius R
$d_1 < 10\text{ mm}$	$d_1 - 1.5\text{ mm}$	2 mm
$10\text{ mm} \leq d_1 \leq 20\text{ mm}$	$d_1 - 2.0\text{ mm}$	2 mm
$20\text{ mm} \leq d_1 \leq 30\text{ mm}$	$d_1 - 2.5\text{ mm}$	3 mm
$30\text{ mm} \leq d_1 \leq 40\text{ mm}$	$d_1 - 3.0\text{ mm}$	3 mm
$40\text{ mm} \leq d_1 \leq 50\text{ mm}$	$d_1 - 3.5\text{ mm}$	4 mm
$50\text{ mm} \leq d_1 \leq 70\text{ mm}$	$d_1 - 4.0\text{ mm}$	4 mm
$70\text{ mm} \leq d_1 \leq 95\text{ mm}$	$d_1 - 4.5\text{ mm}$	5 mm
$95\text{ mm} \leq d_1 \leq 130\text{ mm}$	$d_1 - 5.5\text{ mm}$	6 mm
$130\text{ mm} \leq d_1 \leq 240\text{ mm}$	$d_1 - 7.0\text{ mm}$	8 mm
$240\text{ mm} \leq d_1 \leq 500\text{ mm}$	$d_1 - 11.0\text{ mm}$	12 mm

Hardness

55 HRC or 600 HV,
hardness penetration depth at least 0.3 mm.

Housing bore for the shaft seal

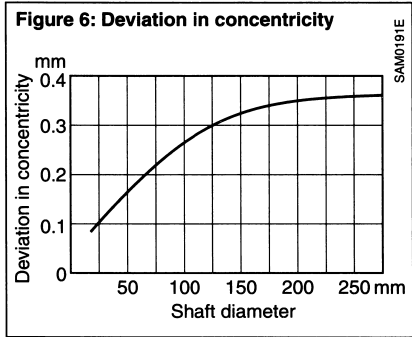
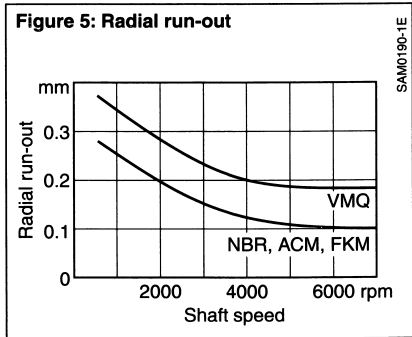
The housing dimensions are shown in Table 4 and Figure 7. The tolerances for the metric sizes comply with DIN 3761 or ISO 6194 so that a good press fit is achieved for a tolerance in the housing bore with ISO H8. For the imperial sizes (inches) the tolerances conform to the American standards. In installation cases where the housing bore has a different tolerance, the seal can be manufactured with a matching interference. A special fit between the seal and the bore may be required for housing bores of soft material, e.g., light alloy, in common with bearing housings with thin walls. The tolerances for seal and bore must in such cases be defined by practical tests.

If a part, e.g., a bearing, is pressed by the seal seat, this can be damaged. To avoid such damage, the seal must be chosen with a larger outside diameter than that of the bearing.

Surface roughness of the housing bore

The values for the surface roughness in the housing bore are specified in ISO 6194-1. The following values are recommended:

$R_a = 1.6$ to $3.2\text{ }\mu\text{m}$,
 $R_z = 6.3$ to $12.5\text{ }\mu\text{m}$,
 $R_{\text{max}} (= R_t) = 16$ to $25\text{ }\mu\text{m}$.



For metal-caged seals (not rubberized) or required gas tightness, a good scoring-free, flat-finish surface quality is required. If the radial shaft seal is glued in the housing, it must be ensured that no adhesive comes into contact with the sealing lip or the shaft.

Installation instructions

The following points must be observed in relation to installing radial shaft seals:

- The installation spaces must be cleaned prior to installation.
- In the case of rubber seals, shafts and seal must be greased or oiled.
- Sharp-edged transitions must be chamfered, rounded or covered.
- When pressing in, it is important to ensure that the seal is not tilted.
- The press-in force must be applied as closely as possible to the outside diameter.
- The seal must be seated centrally and perpendicularly to the shaft.
- The end face of the locating bore is usually used as the stop face; the seal can also be secured with a shoulder or a spacer.

Table 4: Housing dimensions

Width of seal b mm	b_1 min. ($0.85 b$) mm	b_2 min. ($b + 0.3$) mm	r_2 max. mm
7	5.95	7.3	0.5
8	6.80	8.3	
10	8.50	10.3	
12	10.30	12.3	
15	12.75	15.3	0.7
20	17.00	20.3	

Figure 7: Housing dimensions: installation depth and lead-in chamfer
Quantities, see Table 4.

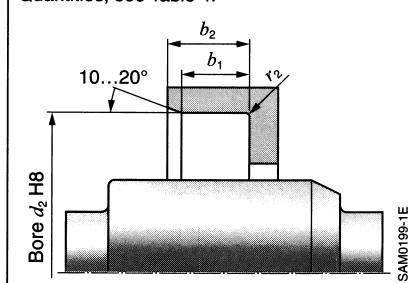


Figure 8 shows different press-in situations for the radial shaft seal with suitable mounting tools or devices.

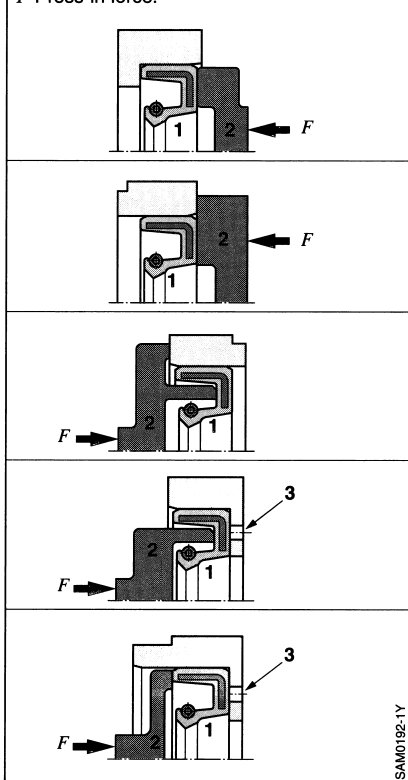
Removal and replacement

Generally speaking, the removal of seals does not pose any difficulties. A screwdriver or similar is usually sufficient to remove them. The seal is damaged in this process. After a machine has been repaired or overhauled, new radial shaft seals should always be fitted, even if the old seals still appear to be intact.

The sealing edge of the new ring should not rest on the old running point. This can be achieved by

Figure 8: Installation aids in the mounting of radial shaft seals

- 1 Radial shaft seal,
 - 2 Mounting tool,
 - 3 Removal hole.
- F Press-in force.



- replacing the shaft protective sleeve,
- pressing into the locating bore to a different depth,
- reworking the shaft and fitting a shaft protective sleeve.

Usage parameters

Overpressure

If the sealing cup is subjected to overpressure, it is pressed against the shaft, increasing the contact area of the sealing lip against the shaft. This increases friction and heat build-up. At overpressure, therefore, the guide values for the maximum permissible circumferential speed are not applicable, but instead these must be decreased in proportion to the intensity of the pressure. At high circumferential speeds, however, already overpressures of 0.01 to 0.02 MPa can cause problems. By using an additional back-up ring, it is possible to use the DIN types A, AS, C and CS pictured in Figure 2 above 0.05 MPa. The separate back-up ring is to be adapted to the back face of the seal-

ing cup. It should not however rest on the sealing cup while there is no overpressure applied. The back-up ring must be adjusted to fit as accurately as possible.

In type TRU the strengthening ring is fashioned in such a way that it supports the sealing cup (Figure 9). Type TRP has a short and sturdy sealing lip which permits an overpressure without additional support. When a back-up ring is installed or when thrust rings are used, overpressures of 0.4 to 0.5 MPa can be permitted at moderate circumferential speeds.

At high overpressures seal types with rubber outer casings should be chosen (type TRA) so as to prevent leakage at the locating bore. At overpressure the seal is at risk of shifting in the axial direction in the housing bore (pressing out). This can be avoided by securing the seal with a shoulder, a spacer ring or a retaining ring.

Circumferential speed and rotational speed

Different sealing-cup designs influence the level of friction and thus result in varying temperature increases. The upshot of this is that the different sealing-cup designs permit different circumferential speeds. Figure 10 features guide values for the maximum permissible circumferential speed for sealing elements without protective lip made of NBR, ACM, FKM

Figure 9: Permissible pressure of the medium to be sealed for supported radial shaft seals and for pressure seals

- Diagram,
- Type TRU,
- Type TRP/6CC,
- Type TRA/CB with back-up ring.

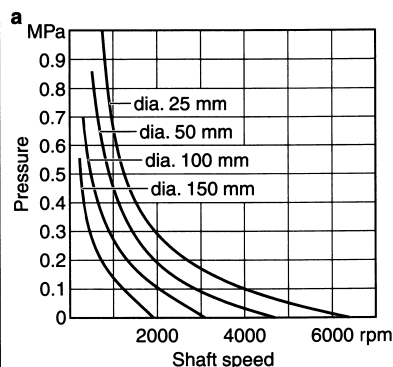
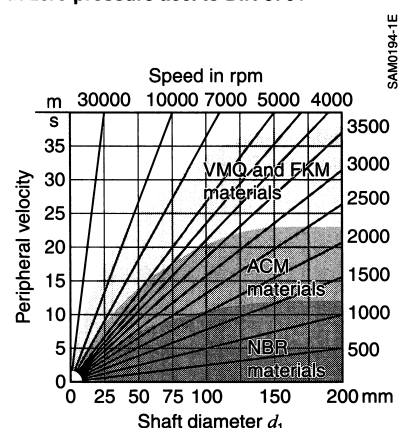


Figure 10: Permissible speeds in state at zero pressure acc. to DIN 3761



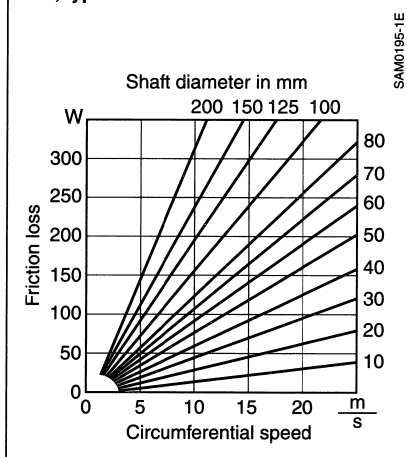
and VMQ in operation at zero pressure and where adequate lubrication or cooling of the sealing edge is guaranteed by the medium to be sealed. The permissible continuous temperatures in Table 2 must be taken into account and may not be exceeded. The bend demonstrates that larger shaft diameters permit higher circumferential speeds than smaller shaft diameters. This is because there is greater heat dissipation as the shaft cross-section increases.

Friction losses

Friction loss is often of a magnitude that must be taken into account. This applies particularly when lower power output is transmitted. Friction loss is influenced by the following factors: Seal design and seal material, spring force, rotational speed, temperature, medium, shaft design, and lubrication.

Figure 11 demonstrates which friction losses in watts are caused by a radial shaft seal without protective lip when it is installed according to the manufacturer's technical directions. In certain cases the friction loss can be reduced by special sealing-lip design. Decreasing the spring force or using a special rubber quality can also achieve the same result. In this connection it must be noted that the friction loss during the seal's "breaking-in period" is greater than that shown in Figure 11. The standard breaking-in period is a few hours.

Figure 11: Friction loss for a nitrile-rubber seal, type TRA



Pneumatic seals

Sealing against compressed air entails a variety of challenges, which are faced with pneumatic valves. Different factors from those relevant to hydraulic seals come to the fore.

As a rule, pneumatic applications are operated with an air pressure of 0.8 to 1.0 MPa, in exceptional cases up to 1.6 MPa, and for high-pressure valves sometimes even up to 5.0 MPa. Speeds are in the range of 0.2 to 1.0 m/s, and in special applications up to 5.0 m/s.

Leakage, friction and lubrication

The leakage of air plays a subordinate role. Friction has a much greater influence on the power output of a pneumatic application. Lubrication is therefore very important. Since the use of oiled air is declining, most pneumatic sealing systems are lubricated with grease on a one-off basis during installation. This lubrication must be maintained over the entire service life. One resulting requirement imposed on a pneumatic seal is hence that the lubricating grease is not wiped in the application. There are therefore differences, when compared with hydraulic seals, in the geometry of the sealing edge and the resulting pressure path.

Alternatively, seals with dry-running properties can be used. However, these are for the most part ruled out for economic reasons.

Sealing system

The most important components of a pneumatic sealing system are pictured in Figure 1.

Piston seals

Piston seals are as a rule "internal" seals and ensure that a pneumatic cylinder can move in that they prevent the medium (air or gas) from overrunning. Piston seals are seated on the moving component and are for the most part double-acting, i.e. the seal acts in both directions. It seals when the piston moves to the left and when it moves to the right. This is often manifested in a symmetrical profile.

Rod seal

Rod seals are as a rule "external" seals and ensure that no medium can escape to the atmosphere. Rod seals are seated in the housing and are for the most part single-acting. The seals acts only in one direction, with these frequently being manifested in an asymmetrical profile.

Dirt wipers

Dirt wipers protect the system against dirt, foreign particles, swarf, and moisture. This prevents premature failure of the components. Single- and double-acting wipers are used, depending on the application and sealing system.

Cushioning seal

Cushioning valves damp the impact of the piston and also assume the function of a non-return valve.

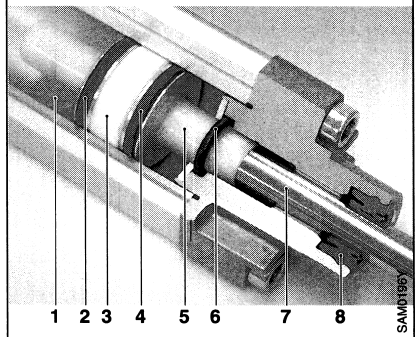
Guide band

Guide bands guide the piston and the rod. They accommodate radial forces, absorb vibrations, and prevent metal-to-metal contact.

Figure 1 shows a standard cylinder comprising the above-mentioned components.

Figure 1: Sealing system of a pneumatic cylinder with piston seal and rod seal

- 1 Cylinder barrel, 2 Piston seal,
- 3 Guide band, 4 Piston seal,
- 5 Piston extension, 6 O-ring, 7 Piston rod,
- 8 Rod seal.



Design types

Aside from classic profile seals such as the O-ring or the X-ring, there are also standard sealing elements for pneumatics. Figure 2 shows some of these seals.

Beyond simple sealing elements the trend is towards the assembly such that complete pistons are already used in many applications (Figure 3).

Materials

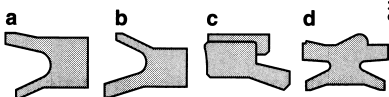
Different materials are used to suit the application: Elastomers, polyurethane and polytetrafluoroethylene (PTFE).

Elastomer seals

Elastomer materials (above all NBR and FKM) are, thanks to their cost effectiveness, the main material used in the production of pneumatic seals. Simply fitting is a positive side effect. Challenges for elastomer seals can be presented by high abrasion.

Figure 2: Pneumatic seal profiles for standard cylinders

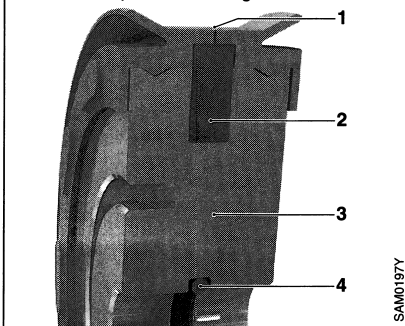
- a) Rod seal, single-acting,
- b) Piston seal, single-acting,
- c) Cushioning seal,
- d) Rod-seal wiper combination
(cf. Figure 1, Profile of rod seal).



SAM0198Y

Figure 3: Pneumatic complete piston

- 1 Double-acting seal arrangement,
- 2 Magnet ring (for position sensing),
- 3 Aluminum piston, 4 O-ring.



SAM0197Y

Polyurethane seals

As belonging to the family of elastomers, polyurethane must be considered separately. As an abrasion-resistant and flexible material, it is often classified between elastomers such as NBR, FKM and PTFE.

PTFE seals

Compared with elastomers, PTFE is frequently not cost-effective because of the manufacturing process involved. It does however offer advantages in conditions of high pressure, wear and friction when there is no lubrication.

Applications

Pneumatic components are used in practically every branch of industry. They feature most strongly in automation engineering or process automation. Here, pneumatic units are used among others for gripping, turning, separating, suction or driving. These components ensure highly efficient, precise and economical production and assembly systems. Open and closed loop control is performed by pneumatic cylinders and valves.

Another area of application is commercial vehicles (trucks, buses, trains). Here they are used in the drive, chassis and suspension, and transmission, in among others non-return valves, pressure-control valves, and switching valves. Cylinders are used in these sectors for example for the precise control of throttle valves.

But pneumatic systems are being used with increasing regularity in other sectors such as biosciences, the food and drinks industries, the chemical process industry, and the electronics industry.

Two-component seals

Multi-component injection molding

Basically, the term multi-component injection molding covers all materials and their combinations which can be combined in the injection-molding process. The following material combinations are the most frequently encountered at present:

- thermoplastic with thermoplastic,
- thermoplastic with thermoplastic elastomer,
- thermoplastic with elastomer,
- multi-color injection molding.

Basic principles

There are various process technologies which can be used in the manufacture of multi-component molded parts. All production technologies are based on the same production process, which works with two or more injection units; however, an explicit distinction is made here between the single-stage process and the two-stage process.

Single-stage process

In the single-stage process the pre-molded part is transported by a handling unit from one side to the other station in the mold. The later multi-component molded part is now either ejected by means of a mold movement (e.g., ejector pin) or else removed by the handling unit and deposited on a cavity-specific basis. Only one closing unit (machine) is used here.

Two-stage process

In the two- or multi-stage process one of the components can even be made of metal with engineers currently endeavoring to replace metal parts as much as possible with technical plastics and hence to achieve a weight reduction. The two- or multi-stage process is performed by way of at least two closing units and hence on two autarkic machines or tools. Here too the insertion parts (metal) or pre-molded parts are inserted into the designated cavities of the respective mold and ejected by mold movement or deposited by a handling unit or a robot on a cavity-specific basis.

Process technologies

A distinction can basically be made in the process technologies between processes with separate contours, components or with contours that run into each other or components.

Composite injection molding

In the case of composite injection molding the combinations should exhibit good chemical adhesion to each other, such as for example a seal to technical components.

Assembly injection molding

Here material combinations that exhibit no chemical adhesion to each other are chosen, since otherwise the function of the component (e.g., a joint) is no longer assured.

Sandwich injection molding or co-injection technique

In this process different material properties are combined with each other, e.g., a rigid outer skin with a foamed core; therefore there does not necessarily have to be adhesion.

Multi-color injection molding

Here the same material is processed in different colors in one component (e.g., in multi-color car tail lights).

Fill simulation (mold flow analysis)

In all the processes mentioned and the design of molds, fill simulation is today one of the most important aids in the field of multi-component injection molding. It delivers detailed results and shows how problems can be avoided across the full injection-molding process. It provides useful information on pressure and temperature distribution and on possible positions of air pockets and joint lines in the component to be produced. In this way, potential flaws in molded parts and molds can in a fill simulation already be detected and eliminated in the development phase.

By detecting these potential flaw sources early, it is possible already at an early stage to eliminate the potential cost drivers in multi-component injection-molded parts.

Composite techniques

With the composite techniques there are a number of ways of combining the materials with each other. The correct composite technique depends here on a number of factors: Use of the component, choice of the materials used, mechanical load on the component, and quantities of the component required. According to these criteria the following composite techniques are currently predominant:

- mechanical bonding by undercuts,
- chemical bonding of the two materials,
- chemical bonding by previous priming (bonding agent),
- chemical bonding by surface treatment of the substrate.

Tool techniques

Rotary table

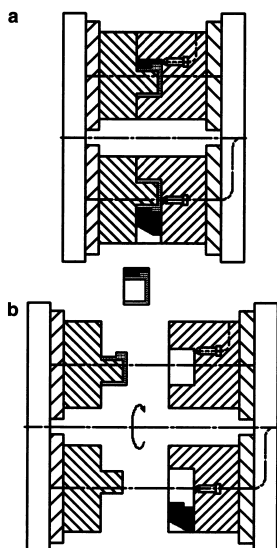
The rotary table is as a rule located between the closing-side mold half and the clamping plate. In this way it is possible to place components from a first cavity into a second cavity. First the material is injected into a first cavity (Figure 1). After intermediate opening and turning, the pre-molded part is now introduced into the second cavity. The cavity additionally opens a cavity for the second plastic. Following a successful injection process the mold is opened again and the finished two-component part is ejected. In parallel with this second operation the next pre-molded part is already produced in the first cavity.

Turning tool

With this turning technique the turning movement is effected not on the machine side, but by a mechanism in the mold (Figure 2). This can be performed for example by a plate in the mold.

Figure 1: Rotary-table process

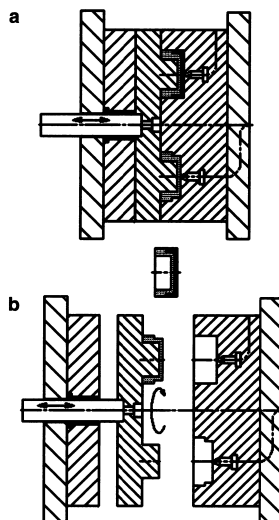
- a) Simultaneous filling of both cavities.
- b) Finished part is ejected, turning of the rotary table on the machine plate including the pre-molded part.



SAM0201-2Y

Figure 2: Turning tool

- a) Simultaneous filling of both cavities.
- b) Finished part is ejected, turning of the mold plate including the pre-molded part.



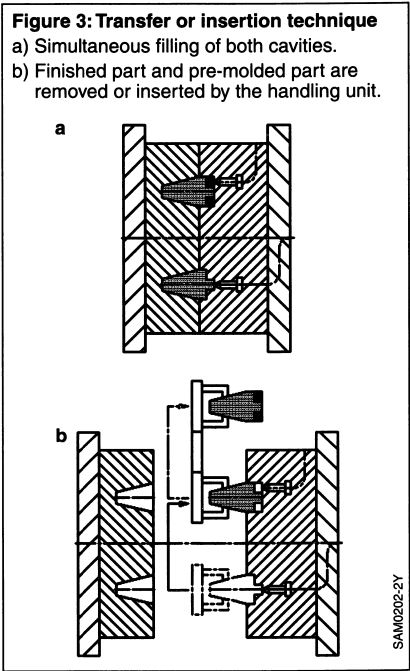
SAM0200-2Y

Turnstiles and turning cores

Instead of turning the complete closing-side mold half, in this solution only an intermediate plate in the mold (also called indexing plate) is turned with the pre-molded part.

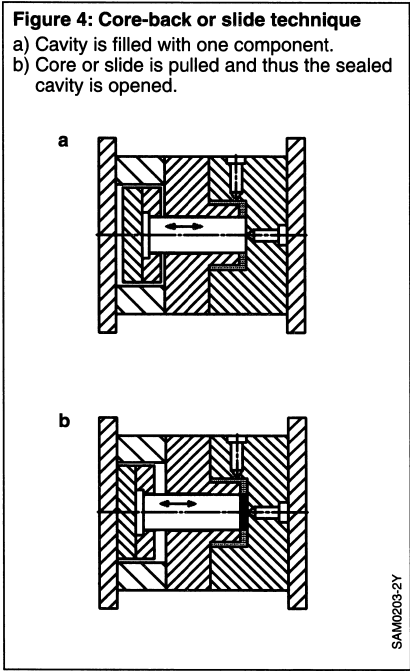
Transfer technique or insertion technique

With the transfer or insertion technique the pre-molded part is transferred or inserted, via a gripper robot or by hand, from one cavity to the other (Figure 3). The second material is then injected here. Transfer technique also refers to the manufacture of the pre-molded part on a machine, removal and insertion in a second machine, and hence a two-stage process.



Core-back technique (slide technique)

The core-back technique (Figure 4) offers the opportunity of manufacturing two-component parts without having to open the mold. The first plastic is injected while the cavity for the second component is kept sealed by a sliding insert in the mold. Once this injection process has finished, the sealed cavity can be opened and the second component processed after the slide is actuated. On contact of the second material the contact-surface temperature can be monitored and adjusted much more effectively with this process than with rotating tool systems.



Rod and piston sealing systems

Hydraulic rod seals

Elements

A rod sealing system as a rule consists of five elements (Figure 1): Rod seal, rod guide, wiper, rod surface, and hydraulic fluid. The individual elements interact with each other. For example, a guide with a lot of play necessitates a seal which can elastically compensate this play. This must be observed particularly in a low-temperature application (-40°C), since the flexibility of the elastomer decreases (see Elastomers). A further interaction can arise for example between the sealing element and the rod. If the rod surface is too rough, wear will occur on the sealing element, possibly on the wiper, and on the guide band. This wear results in leakage, even though the seal is not the cause of this.

Design

Rod seals are asymmetrical seals whose profile is optimized by a sealing edge in such a way that they can as effectively as possible wipe the medium to be sealed from the rod and hence provide a seal. The choice of sealing element is primarily dependent on pressure, required friction, temperature, media resistance, viscosity, lubricity, and price. The following general rule can be formulated: If the price and not the friction at medium pressures is the dominating factor and liquids of normal viscosity and lubricity are involved, elastomer grooved rings or polyurethane (PU) can be used. If, on the other hand, low friction or higher pressures are required, PTFE sealing-edge rings are to be recommended. The same applies in the case of high viscosity or poor lubricity.

PTFE seals are activated by an elastomer energizing element – usually an O-ring – since PTFE has no rubber-elastic properties. It does however have very low friction and high dimensional stability. PTFE sealing-edge rings do not wipe the hydraulic fluid out of the rod-surface topography as well as elastomer and PU seals. They should for this reason exhibit sufficiently good back-pumping behavior, i.e. they should have the capacity to

delivery a lubricant layer of a few molecule layers drawn out during the advance stroke back against the primary pressure into the pressure chamber during the return stroke. Thanks to this behavior, PTFE seals have the advantage of always running on a few molecule layers, which increases their service life. In contrast, PU and elastomers wipe the fluid from the surface in order to be tight. They exhibit only limited back-pumping capacity. On account of the lubricated extension and retraction and the low friction coefficient, the seal is prevented from squeaking or humming much more effectively than PU or elastomer seals.

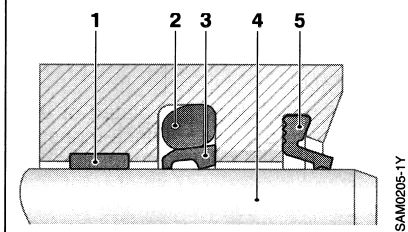
Such systems are used for example in sealing the piston rod of a single-tube shock absorber or in sealing the rack of a hydraulic rack-and-pinion steering system. In this way, even the translatory movements of a positioning cylinder can be sealed in transmissions or in clutches.

Wipers

The function of wipers is to wipe dirt, dust and water from the rod and thus keep them away from the sealing system. Sealing-edge rings are used together with double wipers (with a wiper lip and an inward-wiping sealing lip) or with double-edge wipers (with a lip on which a wiper lip and a sealing edge are positioned). Double-edge wipers have a lower friction than double wipers. Grooved rings manage with a one-lip wiper. They constitute a compromise between a sharp wiper edge with good wiping properties and a

Figure 1: Hydraulic rod seal

1 Guide band, 2 O-ring, 3 Seal, 4 Shaft, 5 Wiper.



minimally rounded edge which does not wipe fluids as consistently.

Elastomers or PU can be used as materials for wipers. In exceptional cases PTFE can be used if substances permanently adhered to the rod are to be wiped. Even for wipers which do not come into contact with the hydraulic fluid, it is important to ensure media resistance on account of the draw-out lubricant layer of the rod seal (this exists for wipers – even PU and elastomers).

Rod guides

Guide rings – actually bands – must be designed in accordance with the requirements with regard to transverse force, friction, media resistance, and temperature. Temperature, friction and media resistance do not pose a problem to the most widely used extended PTFE guides. They cover a temperature range of -40 to $+200$ °C and have the lowest friction values. The media compatibility of PTFE is assured in most cases; only the resistance of the extenders has to be checked. These types of rings and bands can support medium transverse forces. For higher transverse forces a number of rings can be used or a thermoplastic (e.g., PA) can be chosen. It must be noted that all thermoplastic rings significantly lose their load capacity at temperatures in excess of 80 °C. The duroplastic bands familiar from industry are for price reasons only used in exception cases in the automotive sector (see Guides and guide bands).

Hydraulic fluids

Hydraulic fluids are for the most part specified. They should not contain any solid particles. These can reach between seal, guide, wiper or rod surfaces and damage or mark both the elements mentioned and the rod surface; this will then cause leakage. Lubricity and viscosity must be taken into account in the choice of seal.

Rod surface

A hardened or chromium-plated or even otherwise hard-coated surface (consultation with the seal manufacturer here) can at least reduce the problem of striation by hard particles. Lesser damage to PTFE seals normally repairs itself. The required

roughness of the rod is dependent on the seal material and the relative speed. PTFE seals require a smoother surface than PU and elastomer seals. The value $R_z < 2.3$ µm applies to PU and elastomer seals with a maximum speed of less than 0.5 m/s. For PTFE seals this value is between 0.5 µm and 1.3 µm. The low value applies to high speeds and high frequencies (from approximately 0.8 m/s or 5 Hz), the higher value to slower movements. At high pressures and high speeds the surface should have a hardness of HRC > 50 with a case hardness depth $CHD > 0.15$. Generally, closed (plateau-shaped) surface profiles are to be preferred to open (pointed) profiles. All surface values are measured in the seal's axial direction of movement.

Installation of rod seals

Rod seals must be compressed in diameter in the course of installation if no split groove is provided. The compression limit is 13% of the sealing diameter. Compression can be performed with a mounting tool (e.g., cone) with a lead-in chamfer of approximately 10 to 15° which upsets the sealing element to such an extent that it can pass through the bore diameter and snap into the groove.

A PU hose or an axially slotted tube can be used to push the seal into the cone. Subsequent calibration by a rod with a large diameter if possible and a lead-in chamfer of less than 15° are required at least for PTFE seals. The outside diameter of the lead-in chamfer should be smaller than the seal diameter to be calibrated. The calibration is for the most part carried out with the original piston rod, on which a calibrating tip is attached. It is important here that it has a burr-free transition from the tip to the rod which has not sharp edges and thus does not damage the seal. The surface of the calibrating tip or the calibrating tool should have the same surface quality as that of the rod.

Installation by compression (kidney-shaped installation) of PTFE seals, as is sometimes suggested in industrial applications, must be avoided in automotive applications in that permanent deformation of the seals cannot be ruled out. A split groove must be provided if installation by compression is not possible.

Hydraulic piston seals

Seal, guide, medium, surface

The same applies by extension to piston sealing systems as to rod sealing systems. Normally, piston sealing systems have no wipers. This function is at least assumed in part by piston guide rings or piston bands. Piston seals are as a rule not subject to the same high tightness requirements ("only" internal leakage) as rod sealing systems, which must keep oil safely away from the atmosphere. Piston seals are to exhibit equally good tightness in both directions of movement. For this reason, they have a symmetrical cross-section for the most part. Alternatively two grooved rings – back to back – must be used.

One interesting solution is a PTFE piston seal, which activates a sealing edge in each case as a function of the applied pressure. The pressure allows the seal to tilt in such a way that an edge is formed depending on the pressure direction and thereby can operate like a rod seal.

Piston guidance

The same principles as for rod sealing systems and rod guides apply to the choice of the right piston guide rings or guide bands. It proved to be advantageous to place a guide ring or guide band on either side (right and left) of the seal. This enables the guides to work to a limited extent as wipers (see Guides and guide bands).

Hydraulic fluids

See here Hydraulic fluids for rod sealing systems.

Piston surfaces

The same surface-roughness and surface-hardness figures apply to piston seals as to rod sealing systems. A non-hardened, drawn or rolled surface can also be used in special applications. Surface compression produces textures which allow the surface to become harder.

Installation of piston seals

Piston seals can be stretched via a cone with a lead-in chamfer of approximately 10 to 15° onto the piston diameter if the elongation does not exceed 30 % of the inside diameter. The elongated seals then

snaps into the groove. A PU hose or an axially slotted installation tool (e.g., tube) can be used to push the seal into the cone.

Subsequent calibration is required for PTFE seals. This can be done with a separate calibrating tool or with the cylinder itself. The calibration tool is a cylinder with a lead-in chamfer of less than 15° and a bore diameter as close as possible to the piston diameter. The outside diameter of the lead-in chamfer should be at least as large as the outside diameter of the seal to be calibrated. Both in relation to the calibration tool and the lead-in chamfer on the cylinder it is important to ensure that the surface is as good as on the cylinder barrel so as to avoid damage to the sealing element. The transition from the calibration tool to the cylinder should be burr-free and have no sharp edges.

A split groove must be provided if the elongation of the seal exceeds 30 %.

Guides and guide bands

Guide bands have two essential functions to perform in rod and piston sealing systems: Guiding the piston or rod precisely and – if necessary – accommodating transverse forces. Different materials can be used to perform these functions, e.g., bronze, composite materials, or extended PTFE materials. Bronze bushings must be additionally lubricated and are expensive. In the case of composite materials, different guides and guide bands are possible. There are fabric-reinforced duroplastic guide rings, which support a relatively high transverse force, but are not suitable for dry running and even when lubricated have a relatively high friction. The market also offers guides which have a slotted sheet-metal ring on which bronze particles are sprayed and a PTFE layer is rolled in. These bushings offer outstanding guidance quality and low friction, but are at the expensive end of the spectrum.

A further guidance option is provided by wire-mesh bushings which are wrapped in PTFE and likewise have low friction.

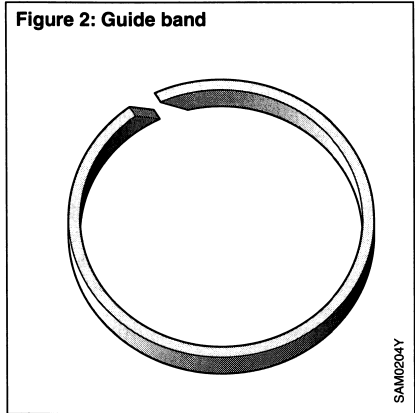
Guide bands made of extended PTFE provide a popular and economical solution in vehicle manufacturing. They exhibit good friction, damp vibrations, and can support transverse forces up to a certain degree.

In the case of guides it must generally be noted that they do not assume the function of a seal. For this reason, the two ends of the guide rings (bands) create a gap (Figure 2). In the case of bushings or guides with narrow play (without a gap), it is important to ensure that no drag streams occur which are superimposed on the system pressure. A bypass may be required in these cases.

The thickness of guide bands depends on the diameter to be guided. The smaller the diameter, the harder it is to bend a thick band for example around a piston and hold it in position. Guide bands with thicknesses ranging from 0.85 to 2.5 mm are used in the automotive sector. The thinner the bands, the smaller their deformations under load and accordingly the better guidance quality they have. Refer to the relevant catalogs for calculation of the permissible load. PTFE materials extended to varying extents are available depending on application and medium.

Typical fields of application for guide bands are, for example, guiding the piston in a single-tube shock absorber or guiding the rod or the piston in a positioning cylinder for translatory movements (transmissions, clutches).

Figure 2: Guide band



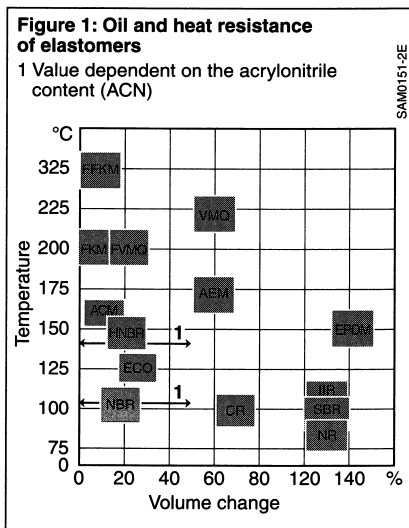
Elastomer materials

General limits of use

The areas of application of elastomer materials are very wide-ranging (see also Materials, Elastomers). Correspondingly different also are the specific requirements which are imposed on the materials in each case (e.g., abrasion resistance, damping performance, insulating effect, long-lasting sealing force, resistance to fatigue cracks, compressive strength, media resistance).

Heat resistance and swelling behavior in oil

The following charts provide a general overview of the different limits of use in accordance with the general classification ASTM D2000 [15], based on the oil and heat resistance of the materials. Figure 1 shows the maximum working temperature and the change in volume after a storage period of 70 hours in the reference oil IRM 903. Figure 2 shows the operating-temperature range. These temperature ranges apply to applications in which contact with media which act aggressively against the respective material is excluded.



Elastomers

Generally speaking, the different elastomers can be roughly characterized as follows. The sequences of capital letters used for the elastomer ID codes are standardized in DIN ISO 1629 (Rubbers) [16].

IIR (isobutene-isoprene rubber)

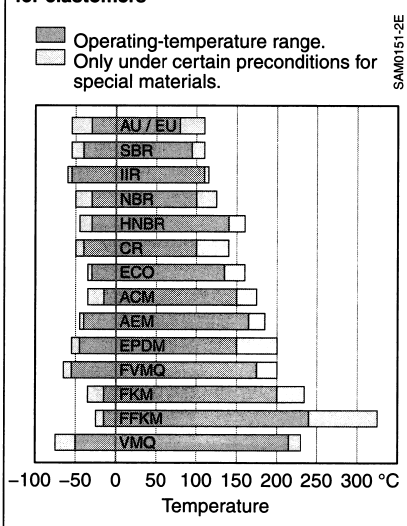
Isobutene-isoprene rubber (butyl rubber) is characterized in particular by its very low permeability to air, water vapor and other gases. In addition, aside from good ozone, weather and aging resistance, IIR also has good resistance to organic and inorganic chemicals. The possible operating temperature is in the range -40 to $+110$ °C (briefly up to $+120$ °C).

Due to the low gas and moisture permeability IIR is used predominantly in tires.

CR (chloroprene rubber)

In general vulcanized chloroprene rubber exhibits relatively good ozone, weather, chemical and aging resistance, in addition high flame retardance, good mechanical properties, good abrasion resistance, and good flexibility at low temperature. The operating-temperature range is -35

Figure 2: Operating-temperature ranges for elastomers



to +90 °C (briefly up to +120 °C). Special types can be used down to -55 °C.

CR materials are used among other things as seals against refrigerant, in outdoor areas, and in adhesive.

SBR (styrene butadiene rubber)

Of particular note in SBR materials are the outstanding mechanical properties, particularly the abrasion resistance. Accordingly, the material is mainly used in tires, but can also be found in V-belts.

In the field of seals SBR plays a lesser role, with resistance to typical media in an automobile such as oil or fuel not being given. The operating-temperature range is -40 to +110 °C (briefly up to +120 °C).

ECO (epichlorohydrin rubber)

The particularity of ECO materials is their good resistance to engine oils and fuels as used in automobiles. Moreover, the extremely low gas permeability, the weather and ozone resistance, and the low compression set must be mentioned.

The operating temperature is specified at -40 to +120 °C (briefly up to +140 °C), special types down to -60 °C.

Because of its outstanding dynamic properties, ECO is used in vehicles particularly for engine bearings and vibration dampers. But ECO is also frequently used in fuel lines.

NBR (nitrile butadiene rubber)

The properties of vulcanized NBR are mainly dependent on the ACN content (ACN: acrylonitrile), which can range between 18 % and 50 %. This determines the oil resistance, which is however always contrary to the flexibility at low temperature.

Vulcanized NBR generally exhibits very good mechanical properties and good abrasion resistance at an operating temperature of -30 to +100 °C (briefly up to +120 °C). Special types can be used down to -50 °C.

NBR is used primarily for seals (also dynamic) or for hoses in contact with mineral oils and greases.

HNBR (hydrogenated nitrile butadiene rubber)

HNBR is produced by selective hydrogenation of the butadiene groups of NBR. In addition to the ACN content the properties are therefore also influenced by the degree of saturation set in the process.

The basic properties of HNBR are comparable with those of NBR. The operating temperature is however clearly improved by the saturation of the main chain and is in the range of -30 to +140 °C (briefly up to +160 °C). Special types can be used down to -40 °C.

They are used mainly in parts which come into contact with mineral oils and greases. However, the material also has good resistance to cooling water and to diesel fuel.

EPDM (ethylene propylene diene rubber)

EPDM has very good heat, ozone and aging resistance, also high elasticity and very good low-temperature performance. The operating temperature is with peroxidic cross-linking in the range -45 to +150 °C (briefly up to +175 °C). With sulfur cross-linking the heat resistance reduces to +130 °C (briefly up to +150 °C).

EPDM is frequently used in glycol-based brake fluids, in cooling water, in refrigerant, and, due to its outstanding weather resistance, for example in door seals.

ACM (polyacrylic rubber)

ACM exhibits very good ozone, weather and hot-air resistance, but only medium strength, low elasticity and relatively unfavorable low-temperature performance. Its operating-temperature range is -20 to +150 °C (briefly up to +175 °C). Special types can be used down to -35 °C.

ACM materials are used as seal materials primarily because of their particular resistance to lubricating oils with high additive levels (also sulfurous) in higher-temperature applications.

AEM (ethylene acrylate rubber)

Compared with ACM, AEM usually exhibits better mechanical properties, much better temperature resistance, and a very good compression set. Its operating-tem-

perature range is -40 to $+160$ °C (briefly up to $+190$ °C).

AEM materials are, like ACM, used as static seals in the underhood area (engine, transmission) primarily because of their particular resistance to lubricating oils with high additive levels in higher-temperature applications.

FKM (fluorine rubber)

Depending on their structure and fluorine content, fluorinated elastomers differ in their media resistance and flexibility at low temperature. Aside from their very good resistance to fuels, oils and other aggressive media, they are characterized by low gas permeability and very good ozone, weather and aging resistance. The operating-temperature range of fluorine rubbers is -20 to $+200$ °C (briefly up to $+230$ °C). Special types can be used down to below -40 °C.

FKM is used, because of its very good resistance to fuels, predominantly in applications such as fuel injectors, fuel lines, and the like.

It is important to mention the really high price of this elastomer group, which varies markedly depending on the polymer type used.

VMQ (silicone rubber)

The particularity of silicone rubbers is their high thermal resistance, the outstanding flexibility at low temperature, good dielectric properties, and above all good weather resistance. Silicone rubbers are hydrophobic and are used above all in the field of aqueous media. Special formulations are resistant to engine and transmission oils and water vapor.

Depending on the version, the possible operating temperatures are in the range -50 to $+200$ °C (briefly in part also up to $+230$ °C).

Silicones are used in automobiles to seal electrical plug-in connections and sometimes also in radiators at high temperatures.

A subgroup, LSR (liquid silicone rubber), is very interesting because of its workability, above all for two-component parts with plastics.

FVMQ (fluorosilicone rubber)

Fluorosilicone rubber has, in comparison with silicones, much better chemical resistance in hydrocarbons, aromatic mineral oils and fuels. The possible operating-temperature range is specified at -55 to $+175$ °C (briefly in part also up to $+200$ °C).

PUR (polyurethanes)

The group of polyurethanes is extremely multilayered. The most diverse ranges of application can thus be individually covered, a standardization of the properties is not possible, and the materials are therefore conceived around their corresponding ranges of application. Depending on the recipe, outstanding recovery performance, extreme flexibility at low temperature, and optimum hydrolysis resistance are achieved. Generally, this material group exhibits low gas permeability, good mineral-oil resistance, high strength and abrasion resistance, but also more unfavorable heat resistance. Depending on the type, temperature operation ranges of -50 to $+110$ °C, briefly even higher, are achievable.

Characteristic data and testing of elastomer materials

Different properties are brought to bear depending on the use to which elastomers are put. To be able to assess the suitability of a material for a certain application, the custom is to conduct tests in advance which are meant to simulate the conditions in the later application. These tests are also intended to define or specify the properties of a material. Some of the most common test methods are described in the following.

Density

Density must be considered as one of the most fundamental properties. This test is – at a constant test temperature – to a large extent not dependent on further parameters such as parts geometry or similar, which is why it tends to be used as a specific material constant as the first identification check in material tests. Usually the density is determined using the buoyancy method according to ISO 1183-1 [17] in which the density is calculated from the measured weight in air and in a test liquid.

Compression set

An important parameter for sealing performance is the compression set of the elastomer material under the influence of temperature. Elastomers demonstrate under load, as well as an elastic component,

a permanent, plastic deformation (Figure 3). The greater the compression-set value, the worse the elastic behavior of a material.

Common test specimens

The compression set is determined according to DIN ISO 815 [18]. A common test specimen is a cylindrical disk with a diameter of 13 mm and a height of 6.3 mm. The test can basically also be carried out with a suitable cross-section on seals (e.g., on O-rings).

Common parameters

The deformation is 25 %. The test temperature is chosen on the basis of the elastomer type. The test duration must be defined and is usually 24 h, 72 h, 168 h, or a multiple thereof.

To be able to compare the results, it is absolutely essential to define all the parameters exactly.

Calculation

The compression set is calculated as follows:

$$\text{Compression set} = \frac{h_0 - h_2}{h_0 - h_1}$$

h_0 Original height,

h_1 Height in the deformed state,

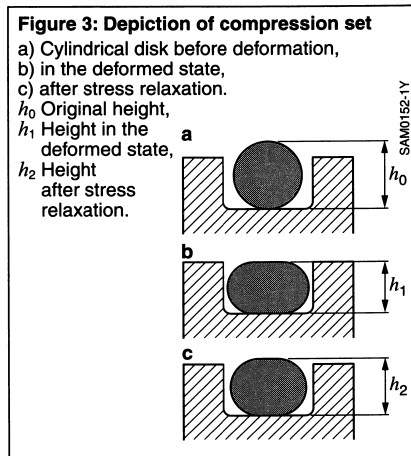
h_2 Height after stress relaxation.

Compressive-stress relaxation

This test, for example according to 3384 [19], is similar to the compression-set test, but here it is not the recovery performance after stress relaxation that is measured but the specifically created counterforce in the compressed state that is determined. It requires much more technical apparatus to be measured, but has a much better correlation with the sealing force for static seals in the installed state.

Common test specimens

- Cylindrical disk with a diameter of 13 mm and a height of 6.3 mm
- O-rings with a diameter of 14 mm and a cross-section of 2.65 mm
- Plain compression rings with a diameter of 15 mm and a square cross-section of 2 mm edge length.



The test can basically be carried out with a suitable cross-section on any seal.

Common parameters

The deformation is 25 %. The test temperature is chosen on the basis of the elastomer type. The test duration must be defined and is usually 24 h, 72 h, 168 h, or a multiple thereof.

Test methods

Two different methods may be used: In method A the deformation and all the counterforce measurements are made at test temperature. In method B the test specimens are stored at the test temperature, with the deformation and all the force measurements being made at room temperature.

The measurement can be conducted in air and in media.

Calculation

The compressive-stress relaxation $R(t)$ is calculated as follows:

$$R(t) = \frac{F_0 - F_t}{F_0} \cdot 100$$

$R(t)$ Compressive-stress relaxation after a specified time t , expressed as % of the initial counterforce,

F_0 Initial counterforce after 30 min,

F_t Counterforce after the specified test duration t .

Hardness

The hardness is one of the most frequently mentioned properties of rubber materials. Nevertheless, the values can be very misleading. Hardness is the resistance of a body to the indentation of an indenting body of a specific shape under a defined compressive force.

Primarily two methods are used for hardness tests on standard test specimens and on finished parts made of elastomer materials:

- Shore hardness (indentation hardness according to the durometer method) according to ISO 7619-1 [20] for measurements on standard test specimens.
- IRHD hardness (International Rubber Hardness Degree) according to ISO 48

[21] for measurements on standard test specimens and finished parts.

The measured values are dependent on

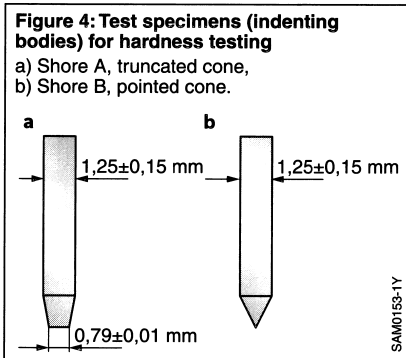
- the modulus of elasticity of the elastomer,
- the visco-elastic properties of the elastomer,
- the thickness of the test specimen,
- the geometry of the indenting body,
- the exerted pressure,
- the speed of the pressure increase,
- the time after which the hardness is registered.

The following also applies here: To be able to compare the results, it is absolutely essential to define all the parameters exactly. Because of these influencing variables, it is moreover not possible to compare the results obtained with a durometer (Shore hardness) directly with IRHD values.

The tests are to be carried out at $23 \pm 2^\circ \text{C}$.

Hardness tests according to Shore A and Shore D

The Shore A hardness tester (truncated cone) can be sensibly used in the hardness range 10 to 90. Harder specimens should be measured with the tester according to Shore D (pointed cone) (Figure 4).



Standard test specimen

- Diameter at least 30 mm
- Thickness at least 6 mm
- Top side and underside smooth and level.

For a thinner material it is permitted to use layers if the minimum specimen thickness is achieved with a maximum of three layers. None of the layers may be less than 2 mm thick.

Hardness tests according to IRHD

Testing the ball indentation hardness according to IRHD (International Rubber Hardness Degree) is applied both to standard test specimens and to finished parts.

The thickness of the test plate must be adapted to the hardness range. According to ISO 48 the hardness is subdivided into two ranges (Figure 5).

Soft: 10 to 35 IRHD

- Specimen thickness above 10 mm bis 12 mm.

Normal: Above 35 IRHD

- Specimen thickness 6 to 10 mm,
- Specimen thickness 1.5 to 2.5 mm (micro-hardness).

Hardness values determined on finished part or specimens of other dimensions deviate as a rule from the values measured on standard specimens. This applies mainly in the case of curved sur-

faces. In order to avoid measuring errors, it is essential to position the measuring tip on curved surfaces at the highest point. In the event of very small cord thicknesses, positioning aids such as magnifying glasses, centering pins and automatic laser positioning can be used.

Testing of tensile stress-strain properties according to DIN 53504 [22] or ISO 37 [23]

This test serves to determine the typical mechanical characteristic values of an elastomer such as tensile strength, elongation at break and the stress values of test specimens of a particular shape when subjected to stress at constant speed up to break (Figure 6).

Tensile strength

The tensile strength R_m is calculated as the quotient from the measured maximum force F_{max} and the initial cross-section A_0 of the test specimen:

$$R_m = \frac{F_{max}}{A_0}$$

Elongation at break

The elongation at break ϵ_R is obtained as the quotient from the change measured at the moment of break $L_R - L_0$ (measured length L_R) and the initial measured length L_0 of the test specimen:

$$\epsilon_R = \frac{L_R - L_0}{L_0}$$

Figure 5: Test specimens (indenting bodies) for hardness testing acc. to IRHD

- Soft, ISO 48 "L"/"CL",
- Normal, ISO 48 "N"/"CNL",
- Normal, ISO 48 "M"/"CM".

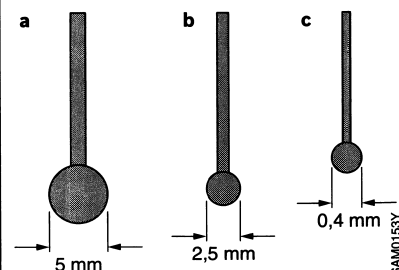
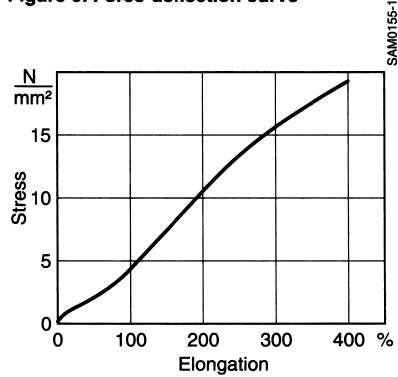


Figure 6: Force-deflection curve



Stress value

The stress value σ is the quotient from tensile force F_i available when a certain elongation is achieved and the initial cross-section A_0 . Usually the value is specified at 100 %.

$$\sigma = \frac{F_i}{A_0}$$

Common test specimens

Common test specimens are S2 tensile bars which are stamped out of a 2 mm test plate (Figure 7).

The test can basically also be carried out on O-rings and is described accordingly. However, the results determined here are not directly comparable with the values of standardized tensile bars.

Common parameters

The tests are generally carried out at the test temperature of 23 ± 2 °C. Often of interest are tests at increased or low temperature which are carried out in special tempering chambers.

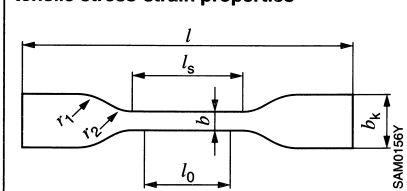
Testing of tear-propagation strength according to DIN ISO 34-1 [24] or ASTM D624 [25]

This test measures the force that is required to tear a specified test specimen either by continuing an already existing cut or through the full width of the test specimen.

Common test specimens

Test specimens are tensile bars of different geometries which are stamped out of a 2 mm test plate. It must be noted that each test specimen leads to a completely different result, which is why only results of the absolutely identical method can be compared with each other.

Figure 7: Test specimen for testing the tensile stress-strain properties

**Test temperature**

The tests are carried out at the test temperature of 23 ± 2 °C.

Testing of media and temperature resistance

Elastomers, like all organic materials, have a limited service life. External factors such as media, temperature, ozone and oxygen influence the material properties. Elastomers can swell, shrink, harden or even crack if they are used under the wrong conditions.

In order to check the suitability of an elastomer under different conditions, it is necessary to conduct compatibility tests according to DIN ISO 1817 [26] in media or according to DIN 53508 [27] in air. Media compatibility can be tested with standardized test medium (e.g., ASTM oils, FAM test fuels) or in concrete operating media from the application.

The influence of a liquid on vulcanized elastomer can cause the following:

- absorption of the liquid by the elastomer,
- extraction of soluble constituents from the elastomer,
- chemical reaction with the elastomer.

The following changes are recorded in this test:

- change of mass, volume and dimensions,
- change of hardness and tensile stress-strain properties according to DIN 53504 [28] after exposure and if necessary after subsequent drying.

Common test specimens

S2 tensile bars according to DIN 53504 are used as test specimens. The test can basically also be carried out on seals (e.g., O-rings).

Common parameters

The test temperature is chosen on the basis of the elastomer type. The test duration must be defined and is usually 24 h, 72 h, 168 h, or a multiple thereof.

To be able to compare results, it is absolutely essential to define all the parameters exactly.

Testing of flexibility at low temperature

The properties of elastomers are very much influenced by temperature extremes. At high temperatures, the material may be subject to anything ranging from hardening to irreversible destruction. On the other hand, the material becomes rigid or hard and brittle at low temperature due to reduced chain mobility and loss of elasticity. This effect is reversible.

Different test methods are used to record this low-temperature effect:

- Calorimetric testing (DSC, Differential Scanning Calorimetry): This enables the T_g value (glass transition temperature) to be determined.
- Static measurement (TMA, Thermomechanical Analysis): This can be used to determine the TR10 value; it describes the recovery of an elongated specimen at low temperature by 10 %, according to ASTM D1329 [29] or ISO 2921 [30].
- Dynamic measurement (DMA, Dynamic-Mechanical Analysis): This can be used to determine the TR value (low-temperature characteristic for the sealing force at low temperature, see above) and further characteristic values.

Each test method leads to a slightly different measuring result, there is no direct correlation between the results.

As well as these analytical methods there is the possibility of determining the freezing-temperature range of a material by means of simple bending tests or a hardness test.

Testing of ozone resistance

Materials which in their polymer structure contain double bonds in the main chain (unsaturated polymers like NBR, HNBR and CR) are susceptible to ozone from the ambient air, above all in the elongated state. Testing of ozone resistance is conducted according to ISO 1431-1 [31] in special test chambers where the test specimens are subjected statically to tensile strain over the test duration. In this test it is assessed

- whether cracks are present in the test specimen after a defined test duration, or
- after what test duration cracks appear with a defined ozone concentration.

Common test specimens

The tests can be carried out on test bars or with a suitable cross-section also on seals.

To be able to compare results, it is essential to define clearly the following parameters: Temperature, ozone concentration, relative humidity, elongation, test duration, and test method.

Testing of rebound resilience

Elastomers are never 100 % elastic; instead, their property is always composed of a viscous component and an elastic component.

The damping property is of interest for certain applications. This can be determined for example by way of the rebound resilience according to DIN 53512 [32] or ISO 4662 [33], whereby a pendulum of defined mass (defined kinetic energy) strikes a test specimen and the height reached after the impact is measured.

Maximum damping corresponds to a value of 0 %. Minimum damping could be at a value of 100 %, but this cannot be achieved.

Common test specimens

The tests are carried out on 12 mm thick test specimens.

References

- [1] ISO 3601: Fluid power systems – O-rings.
 Part 1: Inside diameters, cross-sections, tolerances and designation codes.
 Part 2: Housing dimensions for general applications.
 Part 3: Quality acceptance criteria.
 Part 4: Anti-extrusion rings (back-up rings).
 Part 5: Specification of elastomeric materials for industrial applications.
- [2] DIN 4768: Determination of values of surface roughness parameters R_a , R_z , R_{max} using electrical contact (stylus) instruments; concepts and measuring conditions.
- [3] DIN EN ISO 4287: Geometrical Product Specifications (GPS) – Surface texture: Profile method – Terms, definitions and surface texture parameters.
- [4] http://www.tss.trelleborg.com/de/de/service/design_support/oringcalculator_1/O-ringcalculator.html.
- [5] http://www.tss.trelleborg.com/de/de/service/design_support/chemicalcompatibility_1/chemical-compatibility.html.
- [6] DIN 443: Sealing push-in caps.
- [7] DIN 3761: Rotary shaft lip type seals for vehicles.
 Part 1: Terms; formula symbols, tolerances.
 Part 2: Applications.
 Part 3: Material requirements and methods of test.
 Part 4: Visual irregularities.
 Part 5: Test; measuring requirements and instruments.
- [8] ISO 6194: Rotary shaft lip-type seals incorporating elastomeric sealing elements.
 Part 1: Nominal dimensions and tolerances.
 Part 2: Vocabulary.
 Part 3: Storage, handling and installation.
 Part 4: Performance test procedures.
 Part 5: Identification of visual imperfections.
- [9] DIN 3760: Rotary shaft lip type seals.
- [10] DIN EN 10139: Cold rolled uncoated low carbon steel narrow strip for cold forming – Technical delivery conditions.
- [11] DIN EN 10088: Stainless steels. Part 1: List of stainless steels.
 Part 2: Technical delivery conditions for sheet/plate and strip of corrosion resisting steels for general purposes.
 Part 3: Technical delivery conditions for semi-finished products, bars, rods, wire, sections and bright products of corrosion resisting steels for general purposes.
- [12] VDMA 24317: Fluid power systems – Fire-resistant pressure fluids – Technical minimum requirements.
- [13] DIN 24320: Fire-resistant fluids – Hydraulic fluids of categories HFAE and HFAS – Characteristics and requirements.
- [14] DIN 51524: Pressure fluids – Hydraulic oils.
 Part 1: HL hydraulic oils; minimum requirements.
 Part 2: HLP hydraulic oils; minimum requirements.
 Part 3: HVL hydraulic oils; minimum requirements.

[15] ASTM D 2000: Standard Classification System for Rubber Products in Automotive Applications.

[16] DIN ISO 1629: Rubber and latices – Nomenclature.

[17] ISO 1183-1: Plastics – Methods for determining the density of non-cellular plastics.

Part 1: Immersion method, liquid pycnometer method and titration method.

[18] DIN ISO 815: Elastomers or thermoplastic elastomers – Determination of compression set.

Part 1: At ambient or elevated temperatures.

Part 2: At low temperatures.

[19] ISO 3384: Elastomers or thermoplastic elastomers – Determination of stress relaxation in compression.

Part 1: Testing at constant temperature.

Part 2: Testing with temperature cycling.

[20] ISO 7619-1: Elastomers or thermoplastic elastomers – Determination of indentation hardness.

Part 1: Durometer method (Shore hardness).

[21] ISO 48: Elastomers or thermoplastic elastomers – Determination of hardness (hardness between 10 IRHD and 100 IRHD).

[22] DIN 53504: Testing of rubber and elastomers – Determination of tensile strength at break, tensile strength at yield, elongation at break and stress values in a tensile test.

[23] ISO 37: Elastomers or thermoplastic elastomers – Determination of tensile stress-strain properties.

[24] DIN ISO 34-1: Elastomers or thermoplastic elastomers – Determination of tear strength.

Part 1: Trouser, angle and crescent test pieces.

[25] ASTM D 624: Standard Test Method for Tear Strength of Conventional Vulcanized Rubber and Thermoplastic Elastomers.

[26] DIN ISO 1817: Elastomers or thermoplastic elastomers – Determination of the effect of liquids.

[27] DIN 53508: Testing of rubber and elastomers – Accelerated ageing.

[28] DIN 53504: Testing of rubber and elastomers – Determination of tensile strength at break, tensile strength at yield, elongation at break and stress values in a tensile test.

[29] ASTM D 1329: Standard Test Method for Evaluating Rubber Property Retraction at Lower Temperatures.

[30] ISO 2921: Elastomers – Determination of low-temperature retraction (TR test).

[31] ISO 1431-1: Elastomers or thermoplastic elastomers – Resistance to ozone cracking.

Part 1: Static and dynamic strain testing.
Part 3: Reference and alternative methods for determining the ozone concentration in laboratory test chambers.

[32] DIN 53512: Testing of rubber and elastomers – Determination of rebound resilience (Schob pendulum).

[33] ISO 4662: Rubber – Determination of rebound resilience of vulcanized rubbers.

Detachable connections

Positive or form-closed joints

Operation

Positive or form-closed joints have the task of transferring forces, which themselves maintain contact via mating surfaces, by means of their geometrical shape. The forces are always transferred perpendicularly to the mating surfaces, resulting mainly in compressive and shear stress ([1], [2]).

Positive locking or form closure creates detachable connections, since as a rule a clearance or a transition fit exists (see Tolerances) between the touching surfaces (e.g. between a shaft and the bore of a hub, [3]). Depending on the choice of fit, relative axial movements may occur during operation. If necessary, they must be prevented by suitable locking devices. Among others, retaining rings according to DIN 471 [4] or slotted lock nuts according to DIN 981 [5] are used for this purpose.

Feather-key and Woodruff-key couplings

Feather-key couplings (Figure 1a) are used for the torsion-resistant connection of belt pulleys, gears, coupling hubs, etc. to shafts. Feather keys are sometimes used to secure frictional joints or to fix a specified position in the circumferential direction.

The cheaper Woodruff key (Figure 1b), the round side of which fits into the shaft, is mainly used for this purpose in automo-

Table 1: Symbols and units

Quantity		Unit
D	Diameter	mm
F	Force	N
K_A	Application factor (service factor)	–
M_t	Torque	Nm
b	Width	mm
d	Diameter	mm
h	Height	mm
i	Number of shear surfaces	–
l	Length	mm
l_{tr}	Supporting feather-key length	mm
n	Number of drivers	–
p	Surface pressure	N/mm ²
t_1	Groove depth (shaft)	mm
t_2	Groove depth (hub)	mm
σ_b	Bending stress	N/mm ²
τ_s	Shear stress	N/mm ²
φ	Contact-surface ratio	–

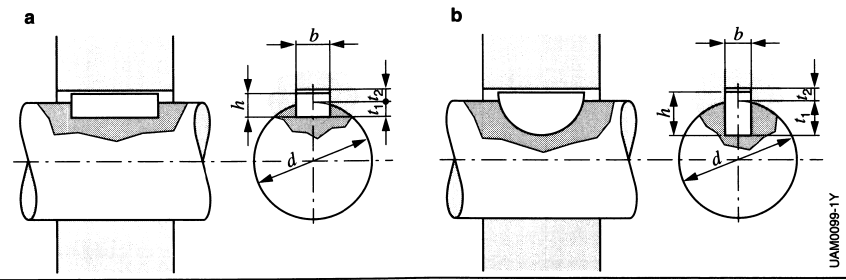
tive construction and to transfer smaller torques.

In the case of feather-key couplings, the groove faces lie against the feather-key faces. In contrast to keyed joints, there is clearance (backlash) between the rear of the feather key and the base of the groove. This means that the forces are transferred exclusively via the flanks of the feather key.

For feather-key width, the tolerance zone $h\ 9$ (key steel to DIN 6880 [6]) is provided. For groove widths b , the tolerance zones specified in Table 2 apply.

Figure 1: Positive or form-closed joints

a) Feather-key coupling, b) Woodruff-key coupling. Symbols, see Table 1.



A sliding fit must be used if a hub must move on the shaft in the longitudinal direction (e.g. gear in manual-shift transmission). Usually, the sliding spring is firmly bolted into the shaft groove. Round (Shape A) and angular (Shape B) feather keys (Figure 2) are manufactured. DIN 6885 [7] defines the standard in terms of their shape and dimensions, depending on the shaft diameter (Table 3).

In practice, feather keys are designed only for surface pressure. If $p \leq p_{\text{perm}}$, the required supporting feather-key length (see Table 1) is

$$l_{\text{tr}} = \frac{2 K_A M_t}{d(h-t_1)n\varphi p_{\text{perm}}}$$

For round-faced feather keys (Shape A), the feather-key length is $l = l_{\text{tr}} + b$. For straight-faced ones (Shape B), it is $l = l_{\text{tr}}$. For the permissible surface pressures, the standard gives $p_{\text{perm}} = 0.9 R_{e, \min}$, where $R_{e, \min}$ is the minimum yield point of the shaft, hub or feather-key material. With one feather key ($n = 1$), the contact-sur-

face ratio is $\varphi = 1$, and with two feather keys, it is $\varphi = 0.75$.

Profiled shaft-hub connections

Instead of inserting multiple feather keys into shaft grooves, the shaft cross-section can also be shaped directly in the form of a polygonal profile, and the mating hub cross-section has a corresponding shape. (Table 4). Profiled shafts are also used if the hub must be axially movable relative to the shaft (e.g. steering column for a height-adjustable steering wheel). The profiled shaft connection has the advantage

Figure 2: Feather-key shapes

a) Shape A, b) Shape B.
Symbols, see Table 1.

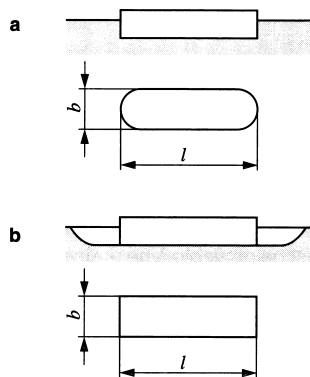


Table 2: Tolerances for groove widths

Groove fit	Fixed fit	Easy fit	Sliding fit
in hub	P 9	N 9	H 8
in shaft	P 9	J 9	D 10

Table 3: Feather-key dimensions according to DIN 6885

Shaft diameter d		Width $\times$ height $b \times h$ mm	Groove depths		Length l mm
over mm	up to mm		t_1 mm	t_2 mm	
6	8	2 $\times$ 2	1.2	1.0	6...20
8	10	3 $\times$ 3	1.8	1.4	6...36
10	12	4 $\times$ 4	2.5	1.8	8...45
12	17	5 $\times$ 5	3.0	2.3	10...56
17	22	6 $\times$ 6	3.5	2.8	14...70
22	30	8 $\times$ 7	4.0	3.3	18...90
30	38	10 $\times$ 8	5.0	3.3	22...110
38	44	12 $\times$ 8	5.0	3.3	28...140
44	50	14 $\times$ 9	5.5	3.8	36...160
50	58	16 $\times$ 10	6.0	4.3	45...180
58	65	18 $\times$ 11	7.0	4.4	50...200
65	75	20 $\times$ 12	7.5	4.9	56...220
75	85	22 $\times$ 14	9.0	5.4	63...250
85	95	25 $\times$ 14	9.0	5.4	70...280
95	110	28 $\times$ 16	10.0	6.4	80...320
Feather-key lengths in mm:		6, 8, 10, 12, 14, 16, 18, 20, 22, 25, 28, 32, 36, 40, 45, 50, 56, 63, 70, 80, 90, 100, 110, 125, 140, 160, 180, 200, 220, 250, 280, 320			

that it does not require any additional separator (feather key) to transfer torque. The hub can be centered on the shaft over the smallest shaft diameter. This is called internal centering and very smooth running can thus be achieved.

The hub can also be centered via the flanks of the drivers or the toothing. This flank centering ensures very low circumferential backlash between shaft and hub and is therefore very suitable for alternating and jerky torques. As with the feather key, a rough design is drawn up based on surface pressure.

Bolt and pin connections

Bolt and pin connections provide a simple and inexpensive means of connecting two or more components. These are among the oldest and most widely used types of connection.

Bolt connections

Bolt connections are mainly used for joining linkages (Table 5), shackles, chain links, and connecting rods, as well as axles for bearing impeller rings, rollers, levers, etc. Since relative movements occur in these connections, at least one part must be movable. The main prevailing stresses are surface pressure (Table 6) and shear. Bending stress is negligible in most cases. It only occurs to a significant extent in the case of bolt connections which are relatively long in relation to their diameter.

Pin connections

Pins are suitable for the permanent connection of hubs, levers, and set collars on shafts or axles (e.g. transverse-pin joint). They also secure the precise position of two machine parts, and are suitable as guide pins to fix springs, etc. (Table 5). Since they are forced into holes as press fits with tolerances, all parts are permanent. The pin can be knocked out with a punch without damaging the components.

Table 4: Profiled shaft-hub connections

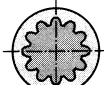
Designation	Standard	Illustration	Driver	Centering	Contact-surface ratio
Spline shaft	ISO 14 [8] DIN 5464 [9]		UAM0101-1Y Prismatic driver	Inner	$\varphi = 0.75$
				Flanks	$\varphi = 0.9$
Toothed shaft with grooved toothing	DIN 5481 [10]		UAM0101-2Y Grooved toothing	Flanks	$\varphi = 0.5$
Toothed shaft with involute toothing	DIN 5480 [11] DIN 5482-1 [12]		UAM0101-3Y Involute toothing	Flanks	$\varphi = 0.75$

Table 5: Bolt and pin connections

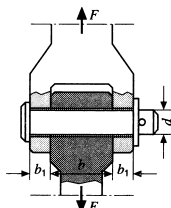
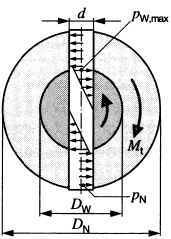
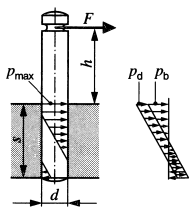
Designation	Illustration	Calculations
Articulated joint	 UAM0102-1a-Y	Surface pressure in fork: $p_G = \frac{F}{2 b_1 d} \leq p_{perm}$ Surface pressure in rod: $p_S = \frac{F}{b d} \leq p_{perm}$ Surface pressure in pin: $\tau_S = \frac{4 F}{i \pi d^2} \leq \tau_{S, perm}$
Transverse-pin joint	 UAM0102-2-Y	Surface pressure in shaft: $p_{W, max} = \frac{6 M_i}{d D_W^2} \leq p_{perm}$ Surface pressure in hub: $p_N = \frac{4 M_i}{d (D_N^2 - D_W^2)} \leq p_{perm}$ Surface pressure in pin: $\tau_S = \frac{4 M_i}{D_W \pi d^2} \leq \tau_{S, perm}$
Guide pin	 UAM0102-3-Y	Maximum pressure: $p_{max} = p_b + p_d = \frac{F}{d s} \left(1 + 6 \frac{h + s/2}{s} \right) \leq p_{perm}$ Bending stress at clamping point: $\sigma_b = \frac{32 F h}{\pi d^3} \leq \sigma_{b, perm}$ Shear stress in clamping point: $\tau_s = \frac{4 F}{\pi d^2} \leq \tau_{S, perm}$

Table 6: Permissible mean surface pressure for bolt and pin connections

Permanent fits			Sliding fits	
Material	Mean surface pressure static	swelling	Material pairing	Mean surface pressure
	p_{perm} N/mm ²	p_{perm} N/mm ²		p_{perm} N/mm ²
Gray cast iron	70	50	St / gray cast iron	5
S 235 (St 37)	85	65	St / CS	7
S 295 (St 50)	120	90	St / Bz	8
S 335 (St 60)	150	105	St hard. / Bz	10
S 369 (St 70)	180	120	St hard. / St hard.	15

Frictional joints

Operation

For frictional joints, press fits are produced in the joints (the friction surfaces being the effective areas) in which the parts to be assembled are in direct contact (Figure 1). The surface pressure p can be generated by bolt forces, keys, elastic separators, or the elasticity of the components themselves. The resulting normal force $F_N = pA$ (with friction surface A) induces a frictional force F_R , which opposes a movement caused by external forces ([1], [2]).

Press fit

Application

For a press fit (cylindrical interference fit), the required surface pressure is produced by the elastic deformation of the shaft and hub, resulting from an interference fit. "Interference fit" is the pairing of cylindrical fitting parts which have an interference before assembly (Figure 2).

Because press fits are easy to produce and can transfer even jerky and variable torques and linear forces, they are suitable for the joints of cylindrical surfaces which do not have to be released afterwards (e.g. gear on shaft, wheel on axle, bushing in housing). Cylindrical press fits are classed as hard-to-detach connections, since with high interference during removal the surfaces of shaft and bores are detrimentally affected.

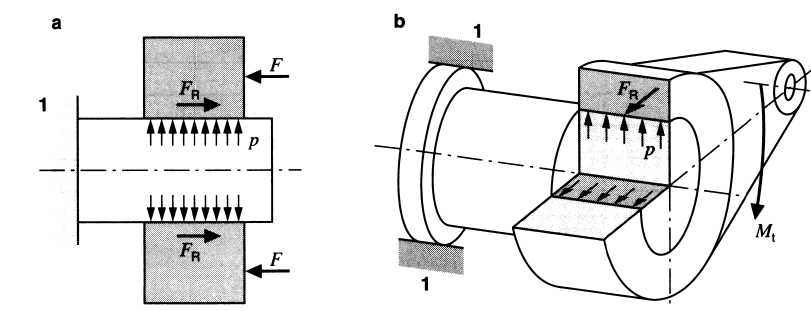
Table 1: Symbols and units.

Quantity		Unit
A	Area (friction surface)	mm ²
C	Taper ratio	—
D	Diameter	mm
E	Modulus of elasticity	N/mm ²
F	Force	N
F_a	Axial force	N
F_N	Normal force	N
F_R	Frictional force	N
K_A	Application factor (service factor)	—
M_t	Torque	Nm
$M_{t, \text{norm}}$	Nominal load torque	Nm
Q	Diameter ratio	—
R_o	Yield point	N/mm ²
R_m	Ultimate strength	N/mm ²
R_z	Surface roughness	mm
S_B	Factor of safety against rupture	—
S_F	Factor of safety against yield	—
U	Tolerance	mm
Z	Allowance	mm
b	Hub width	mm
d	Diameter	mm
l	Taper or lever length	mm
n	Number of bolts	—
p	Surface pressure	N/mm ²
t	Temperature in Celsius	°C
α	Taper angle	°
α_A	Coefficient of linear expansion, outer part	10 ⁻⁶ /K
α_i	Coefficient of linear expansion, inner part	10 ⁻⁶ /K
μ	Coefficient of friction	—
ν	Poisson's ratio	—
ξ	Specific allowance	mm ³ /N
σ_{perm}	Permissible stress	N/mm ²

Figure 1: Frictional joints.

- a) Axially stressed joint,
 - b) Tangentially stressed joint.
- 1 Gripped side of joint.

F Force, F_R Frictional force, p Surface pressure, M_t Torque.



Elastic design of cylindrical interference fits
A press fit must be designed so that at least a minimum surface pressure $p_{\min}$ is present in order to ensure the transfer of the greatest occurring stress, and so that a maximum surface pressure $p_{\max}$ is not exceeded and the components are not overstressed.

Basically, two calculation objectives are possible: The required fit is defined for a given load (Table 2) or the permissible load is determined for a given fit (Table 3). In the following equations index I denotes the inner part (shaft) and index A the outer part (hub) of the press fit.

The diameter ratios (see Figure 2)

$$Q_A = \frac{D_F}{D_{Aa}} \text{ and}$$

$$Q_I = \frac{D_{Ii}}{D_F}$$

and the specific allowance [1]

$$\xi = D_F \left[\frac{1}{E_I} \left(\frac{1+Q_I^2}{1-Q_I^2} - \nu_I \right) + \frac{1}{E_A} \left(\frac{1+Q_A^2}{1-Q_A^2} + \nu_A \right) \right]$$

can be used to design press fits with reference to their function and the required component safety.

The nominal diameter is inserted into the calculation for the joining diameter D_F . When joining an interference, loss results from plastic leveling of the roughness

Table 2: Definition of fits for given stress [1]

Stress M_t or F_a given	$\sigma_{\text{perm}} = R_e / S_F$ or $\sigma_{\text{perm}} = R_m / S_B$ given
Required fit: $p_{\min} = \frac{\sqrt{F_a^2 + \frac{4M_t^2}{D_F^2}}}{\mu \pi D_F b} K_A$	Permissible fit in hub: $p_{\max} = (1 - Q_A^2) \sigma_{\text{perm}} / \sqrt{3}$ Permissible fit in hollow shaft: $p_{\max} = (1 - Q_I^2) \sigma_{\text{perm}} / \sqrt{3}$
Required allowance: $Z_{\min} = p_{\min} \xi$	Permissible allowance: $Z_{\max} = p_{\max} \xi$
Required interference: $U_{\min} = Z_{\min} + 0.8 (R_{zI} + R_{zA})$	Permissible interference: $U_{\max} = Z_{\max} + 0.8 (R_{zI} + R_{zA})$
Select ISO fits with $U_k \geq U_{\min}$ and $U_g \leq U_{\max}$	

Figure 2: Press fits

- a) After joining.
b) Before joining.
1 Inner part (shaft), 2 Outer part (hub).
 F_R Frictional force, p Surface pressure, M_t Torque, b Hub width, D_F Joining diameter,
 D_{Ia} Shaft outside diameter, D_{Ii} Shaft inside diameter,
 D_{Aa} Hub outside diameter, D_{Ai} Hub inside diameter.

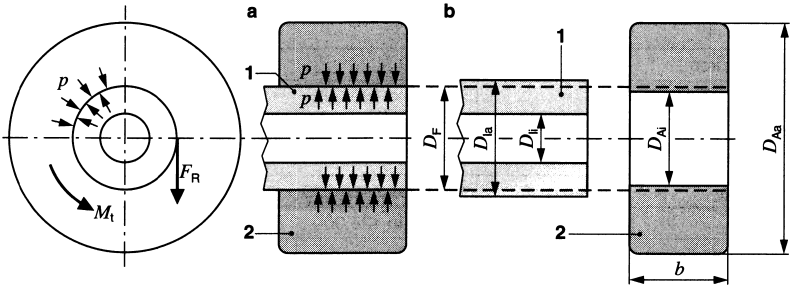


Table 3: Determination of stress for given fit

Smallest interference U_k given	Greatest interference U_g given
Smallest allowance: $Z_k = U_k - 0.8 (R_{z1} + R_{za})$	Greatest allowance: $Z_g = U_g - 0.8 (R_{z1} + R_{za})$
Smallest pressure: $P_k = \frac{Z_k}{\xi}$	Greatest pressure: $P_g = \frac{Z_g}{\xi}$
Permissible stress: $M_t = 0.5 p_k \mu \pi D_F^2 b$ (only M_t) $F_a = p_k \mu \pi D_F b$ (only F_a) $M_t = \frac{D_F}{2} \sqrt{(p_k \mu \pi D_F b)^2 - F_a^2}$ (F_a given) $F_a = \sqrt{p_k \mu \pi D_F b - \frac{4 M_t}{D_F^2}}$ (M_t given)	Hub safety factor: $S_F = \frac{1 - Q_k^2}{\sqrt{3} p_g} R_e$ or $S_B = \frac{1 - Q_k^2}{\sqrt{3} p_g} R_m$ Hollow-shaft safety factor: $S_F = \frac{1 - Q^2}{\sqrt{3} p_g} R_e$ or $S_B = \frac{1 - Q^2}{\sqrt{3} p_g} R_m$

peaks. According to DIN 7190 [13], the smoothing of the two surfaces with 40 % of the averaged peak-to-valley heights R_{z1} (shaft) and R_{za} (bore) is taken into account in the calculation.

The maximum tensions occur according to the equations in Table 2 and Table 3 at the internal diameters of hollow shaft and hub. Solid shafts are noncritical, and do not usually have to be calculated.

Assembly

There are two types of press fit depending on the assembly method: linear and transverse. Linear press fits are produced by “cold” assembly at room temperature. The large press-fit forces required are usually applied by hydrostatic presses. The press-fit velocity should not exceed 2 mm/s. For the press-fit force:

$$F_e = \frac{(U_g - 0.8 (R_{z1} + R_{za})) \mu \pi D_F b}{\xi}.$$

Before transverse press fits are made, either the outer part is expanded by heating, or the diameter of the inner part is reduced by supercooling, so that the parts can be assembled stress-free. If the outer part is heated, it shrinks onto the

Table 4: Poisson ratio, modulus of elasticity, and linear coefficient of thermal expansion for metallic materials

Material	Poisson ratio ν	Modulus of elasticity E [N/mm ²]	Coefficient of linear expansion α [10 ⁻⁶ /K]	
			Heating	Cooling
Gray cast iron	0.24	100,000	10	-8
Malleable cast iron	0.25	90,000 to 100,000		
Steel	0.3	200,000 to 235,000	11	-8.5
Bronze	0.35	110,000 to 125,000	16	-14
Red brass	0.35 to 0.36	110,000 to 125,000	17	-15
CuZn	0.36	80,000 to 125,000	18	-16
MgAl ₈ Zn	0.3	65,000 to 75,000	23	-18
AlMgSi	0.34			

inner part as it cools (shrink fit). If the inner part is cooled so that it expands when it is heated to room temperature, this is an expansion fit. To make force-free assembly possible, a joint clearance of $ED = 0.001 \cdot D_F$ must be provided.

With the ambient temperature t_u , the maximum interference U_g and the coefficient of linear thermal expansion α (Table 4) the assembly temperature of the hub comes out for a shrink fit at

$$t_A = t_u + \frac{U_g + \Delta D}{\alpha_A D_F}$$

and for an expansion fit at

$$t_I = t_u - \frac{U_g + \Delta D}{|\alpha_I| D_F}.$$

Tapered connection

Application

A tapered connection (tapered interference fit) is suitable for transferring dynamic forces and torques. They are used predominantly to secure hubs to shaft ends, e.g. the V-belt pulley on the alternator or the drill head in the drill spindle. Taper and taper angle are standardized, among others also connecting elements

like the terminals on starter batteries, in DIN 254 [14].

Tapered connections for the most part can be retightened and easily detached. As a shaft-hub connection they exhibit no shaft weakening and have very good centering (i.e. no imbalance).

However, the disadvantages are the high production costs and the lack of adjustability in the axial direction. The following applies (Figure 3) to the

$$\text{taper ratio } C = \frac{(d_1 - d_2)}{l}$$

and to the

$$\text{taper angle } \tan \frac{\alpha}{2} = \frac{(d_1 - d_2)}{2l}.$$

The following taper ratios are specified as guidelines (DIN 254 [14], DIN 406 [15]):

$C = 1:5$ connection is easily detachable,
 $C = 1:10$ connection is detachable only with difficulty.

Operation

The effective area pair of a tapered connection takes the form of a truncated cone (frustum). The required surface pressure p is usually applied by an axial bolt force F_a . The relationship between the axial press-fit force F_a and the transferable torque M_t is expressed by the following equation [1]:

$$F_a \geq \frac{2K_A M_{t, \text{nom}}}{\mu_u d_m} \left(\sin \frac{\alpha}{2} + \mu_a \cos \frac{\alpha}{2} \right).$$

Load excesses (impacts) which can occur during operation are taken into consideration with the application factor K_A , which according to DIN 3990 [16] can range between 1 and 2.25. This equation also takes into account the possible difference between the coefficients of friction μ_u in the circumferential direction and μ_a in the axial direction. Self-locking exists if an expulsion force (F_a becomes negative) is required to detach the connection. This means that a torque can be transferred even if the bolt is removed after the parts are axially stressed. In contrast, with a non-self-locking tapered connection, there is no pressure between the effective areas after the axial application force is released. The condition for self-locking is:

$$\frac{\alpha}{2} \leq \arctan \mu_a.$$

Figure 3: Tapered connection

D_a Hub outside diameter,
 d_1 Large taper diameter,
 d_2 Small taper diameter,
 d_m Middle taper diameter,
 l Common length,
 p Surface pressure,
 α Taper angle,
 M_t Torque,
 F_a Assembling force (fastener force).

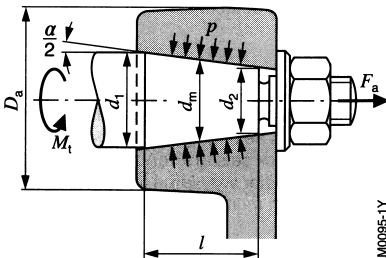
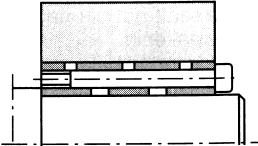
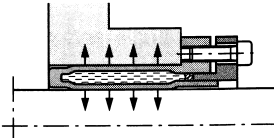
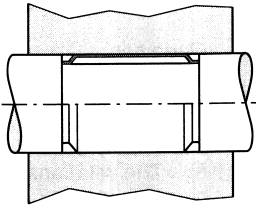
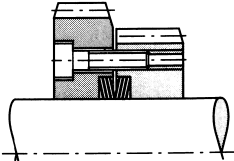
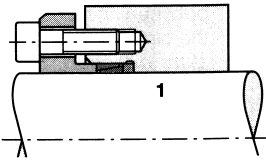
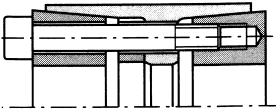


Table 5: Taper-lock joints.

Designation	Illustration	Features
Clamping sleeve (Spieth)	 UAM0096-1Y	Axial distortion is applied to enlarge the outside diameter of the clamping sleeve and reduce the inside diameter. The danger of loosening under dynamic loads is reduced by using long bolts.
Hydraulic hollow-jacket spring collet (Lenze)	 UAM0096-2Y	Axial distortion is applied to generate a pressure in a thin-walled hollow cylinder. The spring collet is self-centering because of even pressure distribution, and thus has true-running properties. At higher temperatures, the thermal expansion of the pressure fluid must be taken into account.
Tolerance ring (Dr Tretter)	 UAM0096-3Y	Tolerance rings are slotted and made from thin wave-shaped sheet metal. The required initial force is provided by forced deformation of the elastic connecting member. They can bridge relatively large machining tolerances, compensate for thermal expansion, and transfer torques.
Star washer (Ringspann)	 UAM0096-4Y	Star washers are thin-walled, very flat conical shells with radial slots. The axial initial force is translated by forced deformation into a fivefold to tenfold radial force. Star washers do not center.
Taper-lock ring (Ringfeder)	1 Precentering  UAM0096-5Y	Taper-lock rings consist of two coaxially arranged conical rings. Radial prestressing is achieved by axially distorting the rings. Taper-lock rings do not center. They are not self-locking and thus can easily be detached.
Taper-lock set (Bikon)	 UAM0096-6Y	Axial distortion is applied by means of the bolts belonging to the taper-lock set. Taper lock sets are very true-running and can transfer very large torques, particularly with multiple pairs of effective areas.

Component safety

The critical component is the hub. It is calculated as an open, thick-walled hollow cylinder. If $Q = d_m/D_a$, the following applies to the safety of the hub, with elastic design, depending on the modified shear-stress theory (DIN 7190):

$$S_F = \frac{1 - Q^2}{\sqrt{3}} \frac{\left(\sin \frac{\alpha}{2} + \mu_A \cos \frac{\alpha}{2} \right) \pi d_m l}{F_{a, \max}}$$

Taper-lock joints

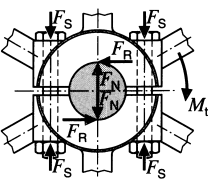
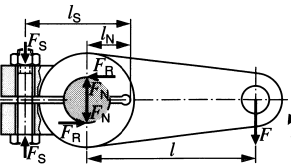
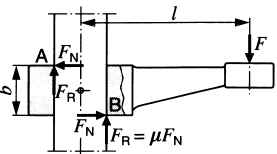
The required surface pressure in the effective areas can also be applied by elastic separators (Table 5). The great advantage of these taper-lock joints is that they can be used to fix hubs, gears, couplings, etc. securely to smooth, cylindrical shafts. Unlike cylindrical interference fits, they can be freely adjusted axially and tangentially, and are above all easily detachable.

They are, therefore, particularly suitable for hubs (e.g. belt pulleys), which must be adjustable and replaceable. Their disadvantages include the space required and the high costs. They are usually designed as specified by the manufacturer (see manufacturer catalogs and Table 5).

Clamp joints

In the case of clamp joints, external forces apply the required surface pressure to the joint, mostly by means of bolts. The joints with a sectional or slotted hub are preferably used for low torques which have little fluctuation. Their advantage is that the hub position is easily adjustable in the axial and tangential directions. They provide a very easy means of fixing wheels or levers to smooth shafts. However, there are also self-locking clamp joints. Here, the tilting force F_K generates edge pressures in A and B (Table 6) to prevent axial movement.

Table 6: Clamp joints

Designation	Illustration	Features
Clamp joint with separated hub		Transferable torque: $M_t = n F_S \mu \frac{\pi}{2} D_F$ n Number of bolts, F_S Fastener force, F_R Frictional force, F_N Normal force, M_t Torque.
Clamp joint with slotted hub		Transferable torque: $M_t = n F_S \mu \frac{\pi}{2} D_F \frac{l_S}{l_N}$ n Number of bolts, F Force, F_S Fastener force, F_R Frictional force, F_N Normal force, M_t Torque, l_S Bolt lever arm, l_N Hub-center lever arm.
Self-locking clamp joint (screw-clamp principle)		Condition for self-locking: $\frac{l}{b} \geq \frac{1}{2\mu}$ F Force, F_R Frictional force, F_N Normal force, μ Coefficient of friction, b Hub width, A, B Contact points.

Keyed joints

Longitudinal keyed joints

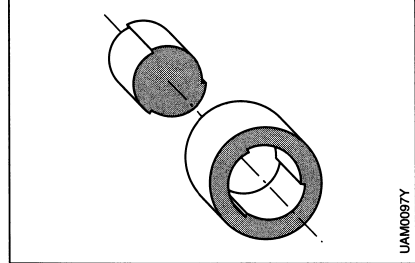
A one-sided radial distortion is achieved by driving in a standardized key (key angle = 0.57°) between the shaft and hub. However, because of their imprecise assembly (hammer assembly) and the resulting eccentricities, keys are only of secondary importance.

Circular keyed joint

A new type of keyed joint is the 3-part circular-key sectional profile (Figure 4). Three circular keys are arranged circumferentially on the cylindrical surface of a shaft (inner part). The hub (outer part) contains an appropriate number of corresponding keys in a cylindrical hole. Twisting results in radial distortion and this can transfer large axial and tangential forces in any direction.

As opposed to press fits, circular keyed joints are detachable. They are used, for example, for shaft-hub connections, for camshafts and for hinges in vehicle construction.

Figure 4: 3-part circular-key sectional profile



Threaded fasteners

Basic principles

Operation

A threaded fastener is used to make secure joints that are detachable any number of times. The purpose of screws and bolts is to stress the mating parts so that the static or dynamic operating forces acting on the joint do not cause any relative movement between the parts ([1], [17], [18], [19], [20]).

When a fastener is tightened or loosened, a screwing motion takes place. In one full rotation of the fastener, an axial shift corresponding to the pitch P occurs.

If a fastener line is uncoiled on a cylinder that has a flank diameter d_2 , it will produce a straight line with a pitch angle $\varphi = \arctan(P/(\pi d_2))$.

In general, fasteners are right-threaded (the straight line rises to the right). Special applications may require left-threaded bolts.

For normal fixing bolts or screws, metric thread profiles (DIN 13 [21], ISO 965 [22]) are used. For pipes, fittings, threaded flanges, etc., pipe threads (DIN ISO 228-1 [23], DIN EN 10226-1 [24]) are used. Refer to Figures 7 through 9 and Tables 7 through 10 for the dimensions of these threads.

Table 1: Symbols and units

Quantity	Unit	Quantity	Unit
A Cross-sectional area	mm ²	d_w Outside diameter of flat fastener-head or nut bearing surface	mm
A_S Stress area	mm ²	f_A Elastic linear expansion by F_A	mm
D_{Km} Effective diameter for friction torque in fastener-head or nut bearing surface	mm	f_{FV} Elastic linear expansion of stressed parts by F_V	mm
E Modulus of elasticity	N/mm ²	f_{SV} Elastic linear expansion of fastener by F_V	mm
F_A Axial operating force	N	f_z Contact stress	mm
F_K Clamping force	N	i Friction-surface pairs	—
F_M Assembly preload	N	l Length	mm
F_N Normal force	N	m Nut height or thread reach	mm
F'_N Normal-force component in plane force polygon	N	n Force-application factor	—
F_{PA} Additional plate force	N	n_S Number of fasteners	—
F_Q Transverse force, operating force applied perpendicular to fastener axis	N	α Flank angle of thread	°
F_U Peripheral force	N	α_a Tightening factor	—
F_S Fastener force	N	μ_G Coefficient of friction in thread	—
F_{SA} Additional fastener force	N	μ_K Coefficient of friction in head bearing surface	—
F_V Preload	N	μ_T Coefficient of friction in parting line	—
F_z Loss of preload due to settling	N	μ'_G Apparent coefficient of friction in thread	—
M_A Tightening torque	Nm	ρ'_G Angle of friction to μ'_G	—
M_G Effective tightening-torque component in thread	Nm	σ_a Alternating cyclic stress on fastener	N/mm ²
M_{KR} Head friction torque	Nm	σ_a Permissible variable stress	N/mm ²
M_L Release torque	Nm	$\sigma_{red,B}$ Equivalent stress in operating state	N/mm ²
P Thread pitch	mm	$\sigma_{red,M}$ Equivalent stress in fitting state	N/mm ²
R_p Yield point	N/mm ²	$\sigma_{z,M}$ Maximum tensile stress in fitting state	N/mm ²
$R_{p0.2}$ 0.2 % yield strength	N/mm ²	σ_z Maximum tensile stress in operating state	N/mm ²
R_p Spring rate of stressed parts (plates)	N/mm	τ_t Maximum torsional stress in thread	N/mm ²
R_S Spring rate of fastener	N/mm	φ Pitch angle	°
R_z Surface roughness	µm	Φ Force ratio	—
W_t Section modulus against torsion	mm ³	Φ_n Force ratio at $n < 1$	—
d' Nominal thread diameter	mm		
d_2 Thread flank diameter	mm		
d_3 Thread root diameter	mm		
d_h Hole diameter of stressed parts	mm		

Calculations for threaded fasteners

The prevailing basis for calculating high-stress threaded fasteners is VDI Directive 2230 [25]. It can be used as a simple, sufficiently precise calculation of the forces and stresses in the thread and the required tightening and loosening torques for a cylindrical single-threaded fastener, which can be considered as a section of a highly rigid multiple-threaded fastener. This means that, in many cases, even complex multiple-threaded fasteners can be considered as single-threaded fasteners. The precondition for this is that the fastener axes are parallel to each other and perpendicular to the parting planes. The components must also be elastic. Another important factor here is that only centrally preloaded and centrally stressed fasteners are considered. For large, eccentric stresses, which may cause the parting line to gape open, please refer to VDI 2230 (Figure 1).

Property classes

According to DIN EN ISO 898-1 [26], the properties of a fastener are identified by two numbers separated by a decimal point. The first number equals 1/100th of the minimum tensile strength; the second is a number that is 10 times the ratio of the yield point in relation to the tensile strength. Multiplying the two numbers gives 1/10 of the minimum yield point (example: $8.8 \rightarrow R_e = R_{p0.2} = 640 \text{ MPa}$).

The property class of a standard nut is identified by one number. This number corresponds to 1/100 of the minimum tensile strength of a bolt of the same property class. To optimize material exploitation, therefore, bolts and nuts of equal property classes should always be paired (e.g. bolt 10.9 with nut 10).

Tightening of threaded fasteners

Prestress

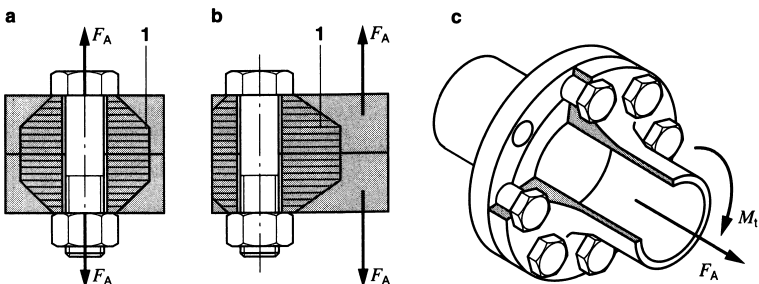
Threaded fasteners are preloaded joints in which the fastener is expanded by f_{SV} by being tightened, and the parts or plates to be tightened are pressed together by f_{PV} . The deformation depends on the dimensions (cross-section and length) and on the materials (moduli of elasticity). According to Hooke's Law, in the elastic range, these are proportional to the prevailing linear force. The ratio of force F to change in length f is the spring rate

$$R = \frac{F}{f} = \frac{E \cdot A}{l}$$

If the rigidities of fasteners and stressed parts are known (they can be calculated according to VDI 2230), the preloaded threaded fastener can be represented by a load-extension diagram. After assembly, an equilibrium of forces arises so that the preload in the bolt and the stressed parts are of identical size (Figure 2a).

Figure 1: Stresses applied to threaded fasteners

- a) Centrally stressed and centrally loaded fastener,
 - b) Eccentrically stressed and eccentrically loaded fastener,
 - c) Multiple-threaded fastener.
- 1 Area of pressure application.



Operating forces

In the case of transversely stressed threaded fasteners (operating force F_O perpendicular to fastener axis), the forces are transferred in the parting plane by friction. Provided that the frictional forces, which are generated by the fastener preload, are greater than the operating forces to be transferred, the assembly load-extension diagram does not change. This means that the fastener "notifies" nothing about the external stress.

At a coefficient of friction μ_T in the parting line, n_S number of fasteners, i number of friction surface pairs, and S_R factor of safety against sliding, the required minimum clamping force is calculated as follows:

$$F_{K,min} = F_V \geq \frac{S_R F_O}{\mu_T n_S i}$$

If an external operating force F_A acts in the direction of the fastener axis according to Figure 1, the bolt is extended by f_A . At the same time the compression of the stressed parts is reduced by the same amount. The bolt is thus subjected to an additional stress of F_{SA} , whereas the stressed parts are relieved by F_{PA} . The additional fastener force F_{SA} thus depends on the rigidity and yield of the fastener (Figure 2b).

The "softer" the fastener (tensile bolt: long and thin), the smaller the additional bolt stress F_{SA} caused by an external axial operating force F_A . This must be the objective, particularly with dynamic operating forces (e.g. for cylinder-head bolts).

Force application

The rigidities of the fastener and the stressed parts also depend on force application. If the external axial operating force F_A is applied directly to the fastener head, the force-application factor $n=1$. For the application of force in the parting line $n=0$. Real force applications range between these two extreme values (Figure 3). In this case, only part of the stressed parts is relieved. This hardens the spring rate R_P of the stressed parts, since grip is reduced. The stressed parts are hammered onto the fastener, which then becomes apparently longer and therefore softer. Small n values therefore result in small additional fastener forces. This has a good impact on the fastener safety factor. However, the clamping force is reduced at the same time, and this is not good for the joint function.

There is no simple method of calculating the force-application factor n , which defines the point of force application. Either n ($0 \leq n \leq 1$) is estimated, or an approximate calculation is made according to VDI 2230. The forces specified in the load-extension diagram can be calculated in accordance with Table 2.

Fastener forces and torques**Calculation model**

The simplest way to represent force ratios in a threaded fastener is by concentrating the surface pressure distributed to all thread turns on a single nut element. During tightening and loosening, the nut

Figure 2: Tightening threaded fasteners

- a) Load-extension diagram on assembly,
b) Load-extension diagram with axial operating force F_A .

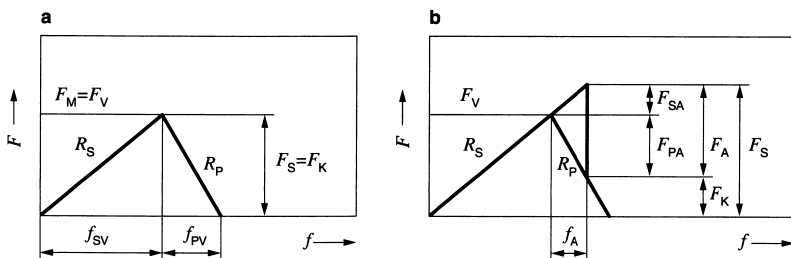


Table 2: Fastener forces (depending on force application)

Force	Force application at fastener head $n = 1$	Force application any $0 < n < 1$	Force application in parting line $n = 0$
Max. fastener force	$F_S = F_V + \Phi F_A$	$F_S = F_V + \Phi_n F_A$	$F_S = F_V$
Clamping force	$F_K = F_V - (1 - \Phi) F_A$	$F_K = F_V - (1 - \Phi_n) F_A$	$F_K = F_S - F_A$
Additional fastener force	$F_{SA} = \Phi F_A$	$F_{SA} = \Phi_n F_A$	$F_{SA} = 0$
Additional plate force	$F_{PA} = (1 - \Phi) F_A$	$F_{PA} = (1 - \Phi_n) F_A$	$F_{PA} = F_A$

Where $\Phi = \frac{R_S}{(R_P + R_S)}$ and $\Phi_n = n \Phi$.

element moves along the bolt thread, which, if unwound, represents a slanting plane or a wedge (Figure 4).

Tightening a threaded fastener

When tightened, the nut element is pushed up the wedge by the peripheral force F_U . The resulting normal force F_N

causes a frictional force F_R which acts in the opposite direction and includes the angle of friction ρ . But since all standardized thread profiles have sloping flanks, as shown in the section in Figure 4, the actual normal force F_N is perpendicular to the sloping thread surface. Therefore only the component $F'_N = F_N \cdot \cos \alpha/2$ appears in the plane force polygon. The following then applies to the frictional force:

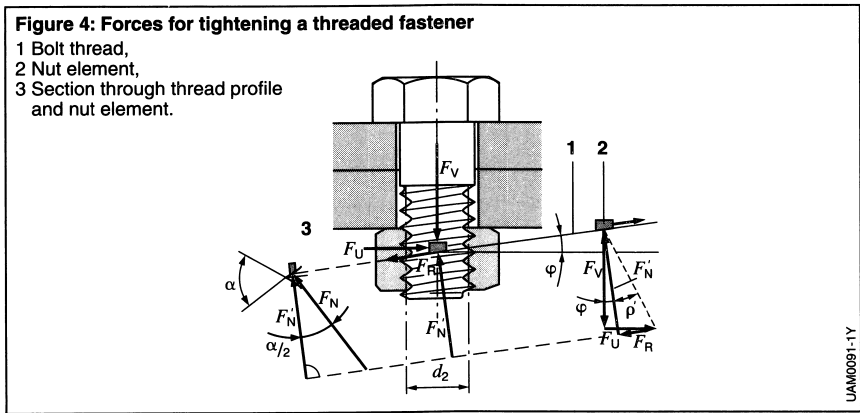
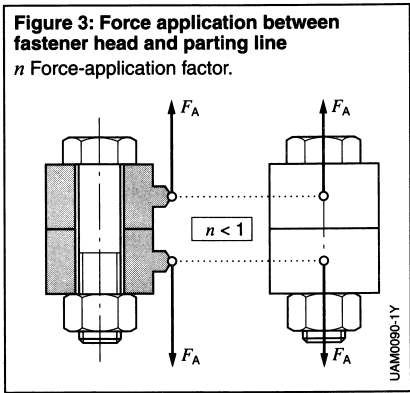
$$F_R = F_N \cdot \mu_G = F'_N \cdot \mu'_G.$$

In order to calculate the force polygon in a plane parallel to the fastener axis, an apparent coefficient of friction is introduced:

$$\mu'_G = \frac{\mu_G}{\cos \alpha/2} = \tan \rho'.$$

When the peripheral force acts on the flank diameter d_2 , the thread torque becomes

$$M_G = F_V \cdot \frac{d_2}{2} \cdot \tan (\varphi + \rho').$$



To tighten a bolt to preload F_V , a head friction torque M_{KR} is required to overcome the friction between the head and nut bearing surfaces, in addition to the thread torque M_G . At a coefficient of friction μ_K and a mean head friction diameter D_{Km} , the head friction torque is calculated as follows:

$$M_{KR} = F_V \cdot \mu_K \cdot \frac{D_{Km}}{2}.$$

The bolt tightening torque applied during assembly is then:

$$M_A = M_G + M_{KR} = F_V \left(\frac{d_2}{2} \tan(\varphi + \rho') + \mu_K \frac{D_{Km}}{2} \right).$$

Loosening a threaded fastener

When a threaded fastener is loosened, the frictional force changes in the opposite direction to the tightening action. The bolt loosening torque (release torque) is calculated as follows:

$$M_L = F_V \left(\frac{d_2}{2} \tan(\varphi + \rho') - \mu_K \frac{D_{Km}}{2} \right).$$

In the case of self-locking threads ($\varphi < \rho'$), the loosening torque becomes negative. This means that a torque must be applied in the opposite direction to the tightening action.

Design of threaded fasteners

Overstressing

If the minimum thread reach $m = (1.0 \text{ to } 1.5) \cdot d$ is maintained, a threaded fastener fails if overstressed, not because the thread turns are stripped, but because the cylindrical screw bolt ruptures.

Assembly stress

When the fastener is tightened to preload F_V , it is stressed up to tensile stress. Due to the thread torque M_G , it is also stressed to torsion. As the friction in the thread prevents the bolt from turning back, the torsional stress also acts after tightening. According to the distortion-energy theory (see Mechanics, Basic principles) the equivalent stress in the bolts is

$$\sigma_{\text{red, M}} = \sqrt{\sigma_{z, M}^2 + 3 \tau_t^2} \leq v \cdot R_{p0.2}$$

with tensile stress

$$\sigma_{z, M} = \frac{F_{V, \max}}{A_S} = \frac{\alpha_A \cdot F_V}{A_S}$$

and torsional stress

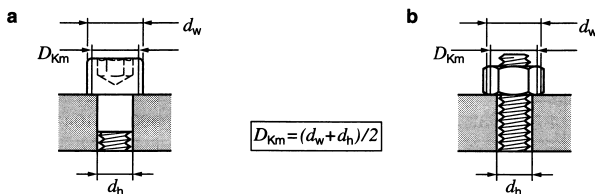
$$\tau_t = \frac{M_{G, \max}}{W_t} = \frac{16 \cdot \alpha_A \cdot F_V \cdot d_2 \cdot \tan(\varphi + \rho')}{2 \cdot \pi \cdot d_3^3}.$$

The tightening factor α_A takes account of the imprecision which is unavoidable during assembly. For torque-controlled tightening (torque wrench) it is $\alpha_A = 1.4$ to 1.6, for pulse-controlled tightening (impact wrench) it is $\alpha_A = 2.5$ to 4.0.

To ensure high functional security, it is necessary to aim for the highest possible material exploitation. This is taken into account by efficiency v . Table 3 shows the permissible preload forces and tightening torques during assembly for different coefficients of friction at 90 % efficiency ($v = 0.9$) of the standardized minimum yield point.

Figure 5: Effective diameter D_{Km} for head friction torque

- a) For Allen screw (d_w = head diameter),
b) For hexagonal bolt/nut (d_w = width across flats).



Static stress

An axial operating force F_A increases tensile stress in the bolt. Since, in the operating state, the effect of torsional stress is less than in the fitted state, VDI 2230 applies to reduced stress:

$$\sigma_{red, B} = \sqrt{\sigma_z^2 + 3(0.5 \cdot \tau_1)^2} < R_{p0.2}$$

at tensile stress:

$$\sigma_z = \frac{F_{S, max}}{A_S} = \frac{\alpha_A \cdot F_V + F_{SA}}{A_S}$$

Vibrational stress

At a dynamic operating force F_A , the variable stress component σ_a must not exceed the permissible variable stress σ_A . The following applies:

$$\sigma_a = \frac{F_{S, max} - F_{S, min}}{2 \cdot A_S} \leq \sigma_A$$

The permissible variable stress σ_A does not depend on the property class, but on the nominal diameter only (Table 4).

Surface pressure between head and nut bearing surface

In the case of a large preload, the surface pressure on the head and the nut bearing surfaces must be checked. Excessive surface pressures can cause plastic deformation and loss of preload. This can result in threaded fasteners coming loose.

The surface pressure p resulting from the maximum fastener force must not exceed the permissible surface pressure limit p_G (guideline values for p_G , see Table 5):

$$p = \frac{4 F_{S, max}}{\pi (d_w^2 - d_h^2)} \leq p_G$$

Table 3: Permissible preloads and tightening torques for regular bolts (acc. to DIN 2230)

Thread	Assembly preload F_V for different coefficients of friction μ in thread						Tightening torque M_A with thread friction $\mu = 0.12$					
	F_V in $10^3 \cdot N$ for $\mu=0.1$			F_V in $10^3 \cdot N$ for $\mu=0.2$			M_A in Nm with $\mu_K=0.1$			M_A in Nm with $\mu_K=0.2$		
	8.8	10.9	12.9	8.8	10.9	12.9	8.8	10.9	12.9	8.8	10.9	12.9
M4	4.5	6.7	7.8	3.9	5.7	6.7	2.6	3.9	4.5	4.1	6.0	7.0
M5	7.4	10.8	12.7	6.4	9.4	11.0	5.2	7.6	8.9	8.1	11.9	14.0
M6	10.4	15.3	17.9	9.0	13.2	15.5	9.0	13.2	15.4	14.1	20.7	24.2
M8	19.1	28.0	32.8	16.5	24.3	28.4	21.6	31.8	37.2	34.3	50.3	58.9
M10	30.3	44.5	52.1	26.3	38.6	45.2	43	63	73	68	100	116
M12	44.1	64.8	75.9	38.3	56.3	65.8	73	108	126	117	172	201
M14	60.6	88.9	104.1	52.6	77.2	90.4	117	172	201	187	274	321
M16	82.9	121.7	142.4	72.2	106.1	124.1	180	264	309	291	428	501
M18	104	149	174	91.0	129	151	259	369	432	415	592	692
M20	134	190	223	116	166	194	363	517	605	588	838	980
M22	166	237	277	145	207	242	495	704	824	808	1151	1347
M24	192	274	320	168	239	279	625	890	1041	1011	1440	1685
M27	252	359	420	220	314	367	915	1304	1526	1498	2134	2497
M30	307	437	511	268	382	447	1246	1775	2077	2931	2893	3386

Threaded-fastener locking devices

The spontaneous loosening of a threaded fastener is caused by complete or partial loss of preload, which in turn is caused by settling events (loosening) or relative movements in the parting line (unscrewing).

Loosening

Contact stress f_z caused by plastic deformation results in loss of preload

$$F_z = \frac{R_p \cdot R_S}{R_p + R_S} \cdot f_z$$

Contact stress f_z in μm depends on the surface properties and the number of parting lines (Table 6).

Total contact stress equals the sum of individual component parts. However, contact stress determined in this way is only valid if the surface pressure limits are not exceeded. Otherwise, significantly greater settling will arise. Locking devices are intended to reduce or compensate for settling.

The following measures provide reliable protection against loosening:

- High preload,
- Elastic threaded fasteners,
- Low surface pressures due to large bearing surfaces and sufficient thread reach,
- Low number of parting lines,
- No plastic or quasi-elastic elements (e.g. seals) with stress.

Unscrewing

Dynamic stresses, particularly perpendicular to the fastener axis, may cause threaded fasteners to come loose despite sufficient preload. If transverse movements can occur, locking devices to prevent unscrewing insure that the joint retains its function. Suitable measures include:

- Avoid transverse movements by positive locking in the parting line,
- Elastic threaded fasteners
- Large grip lengths,
- High preload,
- Suitable locking devices (locking or bonding elements, Figure 6).

Table 4: Permissible variable stress σ_A

Diameter range	M6 to M8	M10 to M18	M20 to M30
Permissible variable stress σ_A in N/mm^2	60	50	40

Table 5: Standard values for surface pressure limit p_G as per VDI 2230

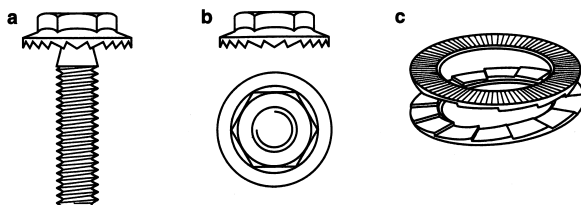
Material	Surface pressure limit p_G in N/mm^2
GD-AlSi 9 Cu 3	290
S 235 J	490
E 295	710
EN-GJL-250	850
34 CrNiMo 6	1,080

Table 6: Contact stress f_z depending on surface properties and number of parting lines

Surface	Load	Contact stress f_z in μm		
		in thread	per head or nut bearing surface	per parting line
$R_z < 10$	Tension/Compression Shear	3.0	2.5	1.5
		3.0	3.0	2.0
$10 \leq R_z < 40$	Tension/Compression Shear	3.0	3.0	2.0
		3.0	4.5	2.5
$40 \leq R_z < 160$	Tension/Compression Shear	3.0	4.0	3.0
		3.0	6.5	3.5

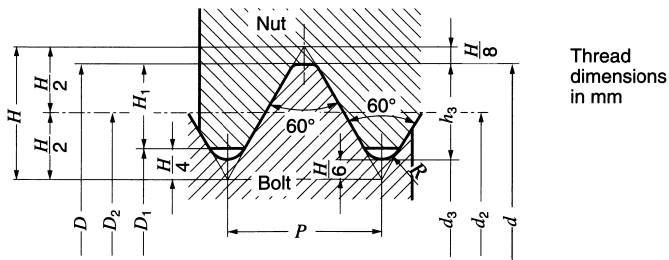
Figure 6: Locking devices (examples)

a) Flat-headed self-locking screw, b) Flat-headed self-locking nut, c) Locking washer pair.



Thread dimensions**Figure 7: Metric ISO thread**

(DIN 13, ISO 965); nominal dimensions.



UAM0041E

Table 7: Metric standard thread

Designation example: M8 (nominal thread diameter 8 mm).

Nominal thread dia. $d=D$	Pitch P	Pitch diameter $d_2=D_2$	Minor diameter d_3 D_1		Thread depth h_3 H_1		Stress area A_s in mm ²
3	0.5	2.675	2.387	2.459	0.307	0.271	5.03
4	0.7	3.545	3.141	3.242	0.429	0.379	8.78
5	0.8	4.480	4.019	4.134	0.491	0.433	14.2
6	1	5.350	4.773	4.917	0.613	0.541	20.1
8	1.25	7.188	6.466	6.647	0.767	0.677	36.6
10	1.5	9.026	8.160	8.376	0.920	0.812	58.0
12	1.75	10.863	9.853	10.106	1.074	0.947	84.3
14	2	12.701	11.546	11.835	1.227	1.083	115
16	2	14.701	13.546	13.835	1.227	1.083	157
20	2.5	18.376	16.933	17.294	1.534	1.353	245
24	3	22.051	20.319	20.752	1.840	1.624	353

Table 8: Metric fine thread

Designation example: M8 x 1 (nominal thread diameter 8 mm and pitch 1 mm).

Nominal thread dia. $d=D$	Pitch P	Pitch diameter $d_2=D_2$	Minor diameter d_3 D_1		Thread depth h_3 H_1		Stress area A_s in mm ²
8	1	7.350	6.773	6.917	0.613	0.541	39.2
10	1.25	9.188	8.466	8.647	0.767	0.677	61.2
10	1	9.350	8.773	8.917	0.613	0.541	64.5
12	1.5	11.026	10.160	10.376	0.920	0.812	88.1
12	1.25	11.188	10.466	10.647	0.767	0.677	92.1
16	1.5	15.026	14.160	14.376	0.920	0.812	167
18	1.5	17.026	16.160	16.376	0.920	0.812	216
20	2	18.701	17.546	17.835	1.227	1.083	258
20	1.5	19.026	18.160	18.376	0.920	0.812	272
22	1.5	21.026	20.160	20.376	0.920	0.812	333
24	2	22.701	21.546	21.835	1.227	1.083	384
24	1.5	23.026	22.160	22.376	0.920	0.812	401

Figure 8: Pipe threads for non self-sealing joints
(DIN ISO 228-1); parallel internal threads and external threads; nominal dimensions

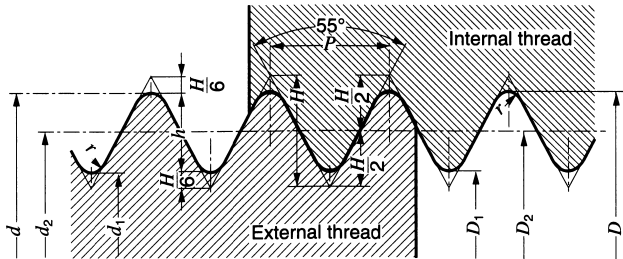
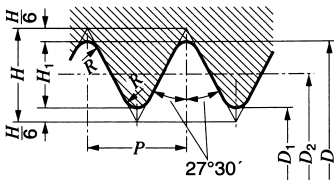


Table 9: Designation example: G1/2 (nominal thread size 1/2)

Nominal thread size	Number of threads per inch	Pitch	Thread depth	Outside diameter	Pitch diameter	Minor diameter
		P mm	h mm	$d=D$ mm	$d_2=D_2$ mm	$d_1=D_1$ mm
$\frac{1}{4}$	19	1.337	0.856	13.157	12.301	11.445
$\frac{3}{8}$	19	1.337	0.856	16.662	15.806	14.950
$\frac{1}{2}$	14	1.814	1.162	20.955	19.793	18.631
$\frac{3}{4}$	14	1.814	1.162	26.441	25.279	24.117
1	11	2.309	1.479	33.249	31.770	30.291

Figure 9: Whitworth pipe threads for threaded pipes and fittings
(DIN 10226-1); parallel internal threads and tapered external threads; nominal dimensions (mm).

Parallel internal thread
(Abbr. Rp)



Tapered external thread
(Abbr. R) (taper 1:16)

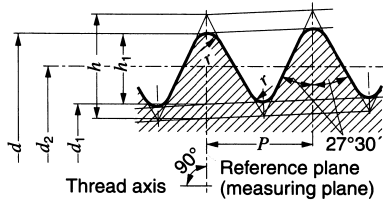


Table 10: Whitworth pipe threads for threaded pipes and fittings

Code		Outside diameter	Pitch diameter	Minor diameter	Pitch	Number of threads per inch
External thread	Internal thread	$d=D$	$d_2=D_2$	$d_1=D_1$	P	Z
R $\frac{1}{4}$	Rp $\frac{1}{4}$	13.157	12.301	11.445	1.337	19
R $\frac{3}{8}$	Rp $\frac{3}{8}$	16.662	15.806	14.950	1.337	19
R $\frac{1}{2}$	Rp $\frac{1}{2}$	20.955	19.793	18.631	1.814	14
R $\frac{3}{4}$	Rp $\frac{3}{4}$	26.441	25.279	24.117	1.814	14
R 1	Rp 1	33.249	31.770	30.291	2.309	11

Areas of application: For joining parallel internal threads to valves and fittings, threaded flanges, etc. with tapered external threads.

Snap-on connections on plastic components

Features

Snap-on connections are a cheap and efficient way of fitting plastic components. They are used to connect housing halves, on plug connectors, and to secure mounting parts in plastic housings. They exploit the high expansibility of the plastics with relatively low rigidity.

All snap-on connections are characterized by the brief excursion of a resilient element in the joining process before it snaps into place behind a locating lug. Depending on the configuration of the joining angles at the snap-on elements, it is possible to produce nondestructively detachable and nondetachable connections (Figure 1).

The basic shapes [27] of snap-on connections are (Table 1):

- Resilient snap-on hooks (bending springs fixed on one side),
- Torsion snap-on hooks,
- Resilient clips,
- Ring-shaped snap-on connections, also segmented (slotted lengthways),
- Spherical snap-on connections, also segmented.

Figure 1: Snap-on connection (principle)

- a) Decisive variables,
b) Joining and release angles (detachable connection: $\alpha_2 < 90^\circ$, nondetachable connection: $\alpha_2 \geq 90^\circ$).
1 Spring element, 2 Locating lug.
 f Spring travel (rear section), l Length, h Thickness at fixation cross-section, F Joining force, Q Excursion force, α_1 Joining angle, α_2 Release angle.

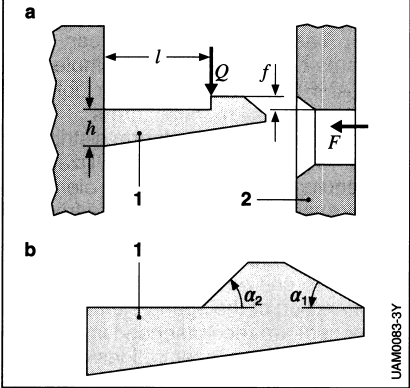
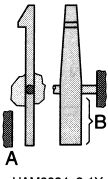
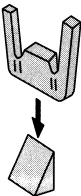
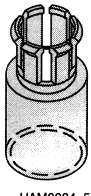
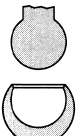


Table 1: Basic shapes and types of snap-on connection [27]

Shape	Hook shape			Ring shape		
				Annular ring, ring groove	Annular ring segmented, ring groove	Hollow-sphere section
Spring element	Bending spring	Torsion spring (+ bending spring)	Connected bending springs with locking element	Annular spring	Annular spring, segmented	Annular spring
Designation	(Bending) snap-on hook	Torsion snap-on hook	Resilient clip	Ring snap-on element	Ring snap-on connection	Spherical snap-on element
Type	 UAM0084_1Y	 UAM0084_2-1Y	 UAM0084_3Y	 UAM0084_4Y	 UAM0084_5Y	 UAM0084_6Y

A Stop against overextension of the torsion axis.

B Button side for releasing the torsion snap-on hook must be more stiffly designed than the side with engagement hook.

Design guidelines and layout

The spring elements are designed to accommodate the permitted elongation of the plastic in the joining process. The least favorable material condition must be taken into account here (e.g. dry polyamide). The elongation-dependent secant modulus

$$E_s = \frac{\sigma_1}{\epsilon_1}$$

is used (determination depicted in Figure 2). The corresponding stress-strain diagrams can be taken for example from the CAMPUS database [29] or obtained from the material manufacturers.

In order to achieve a uniform distribution of strain and optimum material utilization in the bending range of the spring elements, the thickness from the root to the free end should decrease by half. As an alternative, the width to the end of the hook can be reduced to one quarter. Radii at the point of connection of the spring element to the component are recommended in that they help to eliminate concentrations of strain.

When joined, the spring element must be fully returned to its initial state in order to prevent creeping under load and thus permanent deformation. Tensile stress in the snap-on element as a result of operational forces is permitted.

The permitted excursion (spring travel f) in joining is dependent on the geometry of the snap-on hook and the permitted elongation ϵ of the plastic (Table 2). Formulas for different cross-sectional shapes can be taken from the relevant technical literature ([28], [30], [31]) or are provided by special calculation programs.

Included in the calculation of the excursion force Q are the rigidity of the plastic as the secant modulus E_s and the geometry as the bending moment/section modulus W .

The joining force F is calculated from the excursion force Q , the joining angle α_1 (usually 30°), and the coefficient of friction μ between the joining components according to the formula:

$$F = \frac{Q(\mu + \tan \alpha_1)}{(1 - \mu \tan \alpha_1)}.$$

Values for the section modulus W and the friction coefficient μ can also be taken from the technical literature ([28], [30], [31]).

The release force of a snap-on connection is calculated according to the same formula as the joining forced, where the release angle α_2 of the snap-on hook (usually 60°) is to be used. In the case of a nondetachable connection ($\alpha_2 \geq 90^\circ$), the thrust capacity of the snap-on element limits the strength.

Figure 2: Determination of secant modulus E_s

$E_{S1} = \sigma_1 / \epsilon_1$,
 E_0 Modulus of elasticity.

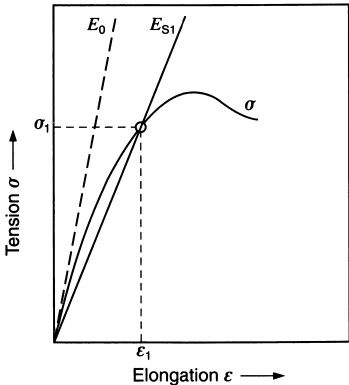


Table 2: Reference values for permitted elongation ϵ for snap-on connections. [28]

(For single, brief joining operation, in case of frequent joining operations reduction of the values by approx. 40 %).

	Material	ϵ
semi-crystalline	PE	0.080
	PP	0.060
	PA (conditioned), POM	0.060
	PA (dry)	0.040
	PBT	0.050
amorphous	PC	0.040
	ABS	0.025
	PVC	0.020
	SAN	0.020
	PS	0.018
glass-fiber-reinforced	PA-GF30 (conditioned)	0.020
	PA-GF30 (dry)	0.015
	PA-GF50 (dry)	0.005
	PBT-GF30	0.015
	PC-GF30	0.018
	ABS-GF30	0.012

Calculation programs

Different plastic manufacturers provide easy-to-use calculation programs as a customer service (e.g. "Snaps" from BASF [32], "FEMsnap tool" from Covestro [33], "FEMSnap" from Lanxess [34]). Most of the material data for manufacturers' product ranges are integrated into these programs. It is also possible and recommended to check the snap-on-hook design by means of FEM analysis within the framework of component development.

References

- [1] H. Haberhauer; F. Bodenstein: *Maschinenelemente*. 18th Edition, Verlag Springer Vieweg, 2017.
- [2] Dubbel, *Taschenbuch für den Maschinenbau*. 24th Ed., Verlag Springer Vieweg, 2014.
- [3] F.G. Kollmann: *Welle-Nabe-Verbindungen*, Springer-Verlag, 1984.
- [4] DIN 471: Circlips (retaining rings) for shafts – Normal type and heavy type.
- [5] DIN 981: Rolling bearing – Locknuts.
- [6] DIN 6880: Bright key steel – dimensions, permissible variations, weights.
- [7] DIN 6885-1: Drive type fastenings without taper action; parallel keys, keyways, deep pattern.
- [8] ISO 14: Straight-sided splines for cylindrical shafts with internal centering; dimensions, tolerances and verification.
- [9] DIN 5464: Straight-sided splines – Heavy series.
- [10] DIN 5481: Serration splines.
- [11] DIN 5480: Involute splines based on reference diameters.
- [12] DIN 5482-1: Internal and external involute spline profiles; dimensions.
- [13] DIN 7190: Interference fits – Calculation and design rules.
- [14] DIN 254: Geometrical product specifications (GPS) – Series of conical tapers and taper angles; values for setting taper angles and setting heights.
- [15] DIN 406-10: Engineering drawing practice; dimensioning; concepts and general principles.
- [16] DIN 3990-1: Calculation of load capacity of cylindrical gears; introduction and general influence factors.
- [17] DIN-Taschenbuch 10: *Mechanische Verbindungselemente 1 – Schrauben*, Nationale Normen. 24th Ed., Beuth-Verlag, 2017.
- [18] DIN-Taschenbuch 45: *Gewinde*. 11th Ed., Beuth-Verlag, 2017.
- [19] DIN-Taschenbuch 140: *Mechanische Verbindungselemente 4 – Muttern*. 10th Ed., Beuth-Verlag, 2016.
- [20] H. Wiegand; K.-H. Kloos; W. Thomala: *Schraubenverbindungen – Grundlagen, Berechnung, Eigenschaften, Handhabung*. 5th Edition, Springer-Verlag, 2007.
- [21]: DIN 13: ISO general purpose metric screw threads.
- [22]: DIN 965: ISO general purpose metric screw threads.
- [23] DIN EN ISO 228: Pipe threads where pressure-tight joints are not made on the threads.
- [24] DIN EN 10226-1: Pipe threads where pressure-tight joints are made on the threads.
- Part 1: Taper external threads and parallel internal threads; dimensions, tolerances and designation.
- [25] VDI Directive 2230: *Systematische Berechnung hochbeanspruchter Schraubenverbindungen* (Systematic calculation of high-stress threaded fasteners). Beuth-Verlag: Sheet 1, 2015; Sheet 2, 2014.
- [26] DIN EN ISO 898-1: Mechanical properties of fasteners made of carbon steel and alloyed steel. Part 1: Bolts, screws and studs with specified property classes – Coarse thread and fine pitch thread.
- [27] U. Delpy u.a.: *Schnappverbindungen aus Kunststoff*. Expert Verlag, 1989.
- [28] A. Maszewski.: *Schnappverbindungen und Federelemente aus Kunststoff*. Anwendungstechnische Information, Bayer AG, 2000.
- [29] <http://www.campusplastics.com>.
- [30] G.W. Ehrenstein (Editor): *Handbuch Kunststoffverbindungstechnik*. Carl Hanser Verlag, 2004.
- [31] T. Brinkmann: *Produktentwicklung mit Kunststoffen*. Carl Hanser Verlag, 2010.
- [32] http://www.plasticsportal.net/wa/plasticsEU-de_DE/portal/show/common/content/technical_resources/calculation_programmes.
- [33] <http://www.plastics.covestro.com/en/Engineering/Tools/FEMSnap-tool>.
- [34] <https://techcenter.lanxess.com/scp/emea/en/techServscp/article.jsp?docId>.

Permanent joints

Welding

Automotive components and subassemblies are joined using a wide and highly varied range of welding and bonding techniques [1], [2]. Resistance and fusion welding are among the most commonly applied welding methods. The overview in Figure 1 shows the most important resistance-welding procedures used in production technology (for procedure types and symbols, see DIN 1910, Part 100, [3]).

Resistance welding

Resistance spot welding

In resistance spot welding, locally applied electrical current is used to melt the contact surfaces of the parts to be joined into a soft or fluid state. The parts are then joined together under pressure (Figure 2a and 2b). The spot-welding electrodes which conduct the welding current also convey the electrode force to the parts to be joined (joining parts). The amount of heat required to create the welding spot is determined in accordance with the equation

$$Q = I^2 R t \text{ (Joule's Law).}$$

The precise amount of heat required Q is a function of welding-current intensity I , resistance R , and welding time t . A good and adequate spot diameter can be achieved by the coordination of welding-current intensity I , electrode force F , and welding time t .

According to the manner in which the current is conducted, a distinction is drawn between bilateral direct resistance spot welding (Figure 2a) and unilateral indirect resistance spot welding (Figure 2b).

The electrode for a specific spot welding operation is selected with reference to shape, outside diameter and point diameter. Since the joining parts should be as free as possible of scaling, oxides, paints, grease, and oils, they receive where necessary appropriate surface pretreatment prior to welding.

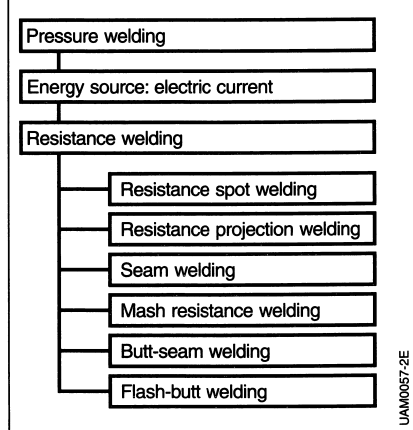
Applications:

- Joining of sheet parts up to an approximate 3 mm single sheet thickness in the lap joint or as a welding flange.
- Multiple-sheet as well as two-sheet joints of different sheet thicknesses and sheet materials.
- Spot-weld bonding in combination with adhesive.

Resistance projection welding

Projection welding (Figure 2c) is a process in which electrodes with a large surface area are employed to conduct the welding current and the electrode force to the workpiece. The projections, which are generally incorporated into the thicker of the joining parts, cause the current to concentrate at the contact points. The electrode force levels the projections partially or completely again during the welding process. A permanent, inseparable joint is produced at the welding seams, i.e. the contact points. One or more than one projection can be welded simultaneously, depending on the type of projection (round, longitudinal or annular) and the power available from the welding appa-

Figure 1: Classification of resistance-welding process in acc. with DIN 1910-100



ratus. Depending on the number of spots welded, a distinction is drawn between single-projection welding and multiple-projection welding.

Resistance projection welding requires high welding currents applied for short periods of time.

Applications:

- Welding parts of different thicknesses.
- Welding multiple projections in a single operation.

Seam welding

In this process, roller electrodes replace the spot-welding electrodes used in resistance spot welding (Figure 2d). Contact between the roller set and the workpiece is limited to an extremely small surface area. The roller electrodes conduct the welding current and apply the electrode force. Their rotation is coordinated with the movement of the part.

Applications:

- Production of sealed welds or seam spot welds (e.g. fuel tanks).

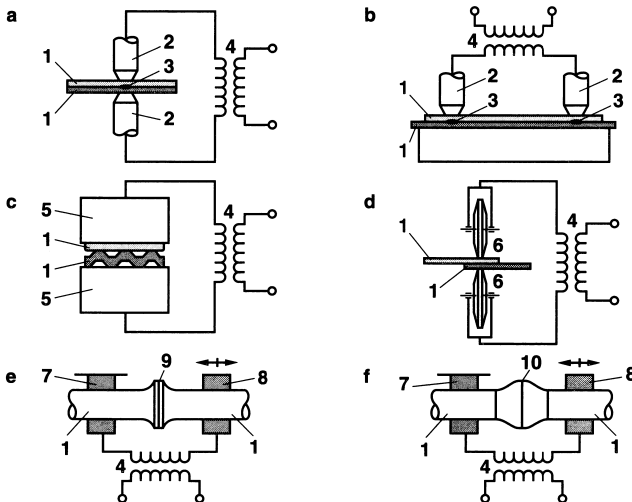
Flash-butt welding

In flash-butt welding, the butt ends of the workpieces are joined under light pressure (Figure 2e) while the flow of current at the contact surfaces produces localized heat and melting (high current density) (current supply via copper jaws). The metal's vapor pressure drives molten material from the contact patches (burn-off). The welded joint is then created by upsetting the workpieces with a high compressive force.

The butt ends should be parallel to each other and at right angles to the direction in which the force is applied (or virtually so). Smooth surfaces are not required. A certain amount of extra length must be factored in to compensate for the losses incurred in the flash-butt welding process.

Figure 2: Resistance-welding processes

- a) Bilateral resistance spot welding, b) Unilateral resistance spot welding,
c) Resistance projection welding, d) Seam welding, e) Flash-butt welding, f) Butt-seam welding.
1 Part to be joined, 2 Spot-welding electrodes, 3 Welding spot, 4 Transformer,
5 Large-surface electrodes, 6 Roller electrodes, 7 Fixed copper jaws, 8
Longitudinally movable copper jaws, 9 Weld joint with seam, 10 Weld joint with bead.



The result is a weld with the characteristic projecting seam.

Applications:

- Joining in the butt joint, e.g. rims, link chains.
- Workshop procedures, e.g. for saw bands.

Butt-seam welding

This process (Figure 2f) employs copper jaws to conduct the welding current to the workpieces to be joined. When the welding temperature is reached, the current switches off. Constant pressure is maintained and the workpieces then weld together (requirement: properly machined butting faces). The process does not completely displace contamination that may be present at the butt ends. The result is a weld with the characteristic projecting bead.

Applications:

- Joining in the butt joint, e.g. shafts, axles.

Fusion welding

The term “fusion welding” describes a process employing limited local application of heat to melt and join the parts. No pressure is applied. Shielded (inert gas) arc welding is a type of fusion welding. The electrical arc extending between the electrode and the workpiece serves as the heat source. Meanwhile, a layer of inert gas shields the arc and the melted area from the atmosphere. The type of electrode is the factor which distinguishes between various techniques:

Tungsten inert-gas welding

In this process, an arc is maintained between the workpiece and a stable, non-melting tungsten electrode. The shielding (inert) gas is argon or helium. The rod-shaped weld metal is supplied from the side (Figure 3).

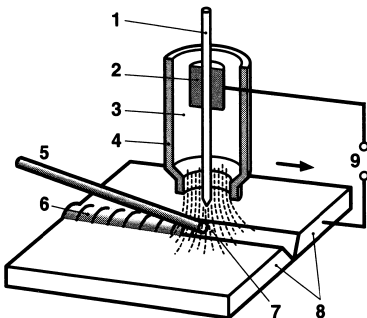
Gas-shielded metal-arc welding

In gas-shielded metal-arc welding, an arc is maintained between the melting end of the wire electrode (material feed) and the workpiece. The welding current flows to the wire electrode via a current-contact nozzle in the welding torch. Metal inert-gas (MIG) welding uses inert gases (slowly reacting and noble gases such as argon, helium, or combinations of the two) as protective gas. MIG welding is used with materials which are particularly sensitive to oxidation, e.g., aluminum, magnesium, titanium, and nickel alloys.

Metal active-gas (MAG) welding, on the other hand, uses an active gas (e.g., CO₂ or mixed gases containing CO₂, argon and sometimes oxygen). MAG welding is used above all for unalloyed and low-alloy steels. Using inert gases with low active-gas admixtures for high-alloy, e.g., stainless steels is also referred to as MAG welding.

Figure 3: Principle of tungsten inert-gas welding

- 1 Tungsten electrode,
- 2 Current contact tube,
- 3 Inert gas,
- 4 Inert-gas nozzle,
- 5 Welding filler,
- 6 Weld seam,
- 7 Arc,
- 8 Workpieces,
- 9 Energy source.



UAM0139Y

Laser-beam welding

Laser-beam welding (or laser welding for short) involves the use of light as the energy source to melt on the workpieces to be welded. The monochromatic laser radiation is generated in a laser source, where the wavelength is determined by the respective excitation medium. CO₂ lasers and NdYAG lasers (solid-state lasers) are used in industrial practice [5], [6]. Newer developments are diode lasers and fiber lasers.

Depending on the wavelength, tube guides with refraction mirrors (CO₂ lasers) or optical fibers (e.g. NdYAG lasers) are required to direct the beam from the beam generator to the weld (Figure 4). To utilize the energy of the laser beam for welding, it is necessary to implement focusing with mirror or lens systems. This allows particularly high energy flux densities to

be achieved at the weld, resulting in the deep-welding effect, which provides for particularly deep, but at the same time, narrow seams. In the simplest case welding is performed without filler metal.

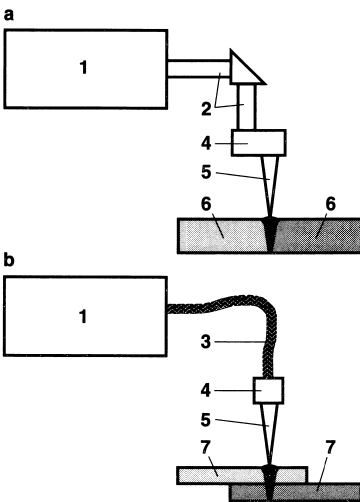
Feeding the workpiece relative to the focusing optics, the optics relative to the workpiece or a combination of the two is used to direct the beam along the seam. Systems with optical fibers are particularly well suited to moving the focusing optics, e.g. during robot-controlled welding of three-dimensional seam geometries.

In the case of remote welding the beam is directed over the workpiece from a relatively large distance (long focal length of the focusing optics) by moving the focusing mirror or the focusing lens inside the focusing optics.

It is possible to operate several machining stations with one beam generator by using beam switches.

Figure 4: Principle of laser-beam welding

- a) Beam direction by mirror,
- b) Beam direction by optical fiber.
- 1 Beam generator,
- 2 Beam direction by refraction mirror,
- 3 Beam direction by optical fiber,
- 4 Focusing optics,
- 5 Focused laser beam,
- 6 Workpieces (butt joint) with weld seam,
- 7 Workpieces (lap joint) with weld seam.

**Applications:**

- Joining of steel sheets using a lap joint in body manufacturing,
- Joining of unalloyed and low-alloy steels as butt joint in chassis and assembly components,
- Joining of high-alloy steels in exhaust-gas systems,
- Joints in seat systems,
- Joining of aluminum alloys (with filler metal).

Other welding processes

The following welding processes are also used in the automotive industry [5], [6]:

- Electron-beam welding,
- Friction welding,
- Arc pressure welding (stud welding),
- Stored-energy welding (pulsed-current arc welding).

Soldering

In soldering, a supplementary material (solder) is melted onto two or more parts of similar or varying metallic composition in order to produce a permanent connection between them. Flux or protective gas is also used [7], [8], [9].

Fluxes (non-metallic substances) are applied to remove oxide layers from the surfaces of the soldering points after cleaning and to prevent new layers from forming: This makes it possible to apply a consistent coat of solder to the joint surfaces. Information on fluxes can be found in DIN EN ISO 9454-1 [10] and DIN EN 1045 [11].

The melting temperature of the solder is below that of the parts being joined. The solder is distributed along the joint to produce the connection without the parts themselves being melted.

The strength of a soldered joint can be equal to that of the base material itself. To achieve the required strength, it is necessary to have narrow soldering gaps in which the adjacent, higher-strength base material prevents the solder from deforming.

Information on the temperature range, the heat source and the geometry of the joint shape (structure) is used to describe the soldered joints. In terms of the working temperature, a distinction is made between soft soldering and hard soldering. The working temperature is defined as the lowest surface temperature at the connection between the workpieces to be joined at which the solder can be melted and distributed to form a bond. Information on solders can be found in the DIN sheets DIN EN ISO 9453 [12], DIN EN ISO 12224-1 [13], DIN EN ISO 17672 [14].

In terms of the joint shapes, a distinction is basically made between close-joint soldering (soldering gap up to 0.5 mm) and V-Joint brazing (soldering gap larger than 0.5 mm).

Distinction by working temperature

Soft soldering

Soft soldering is employed to form permanent solder joints at melting temperatures below 450 °C (as with soldering tin). Soft solders which melt at temperatures of 200 °C and below are also known as quick solders.

Soft-soldering applications:

- Used in electrical engineering for contacts (e.g., flow soldering of printed-board assemblies).

Hard soldering

Hard soldering (brazing) is used to form permanent joints at melting temperatures above 450 °C (as with copper, copper/zinc and silver alloys, e.g. silver brazing filler).

Brazing applications:

- V-joint brazing: steel sheets in body manufacturing, also of different qualities and in the case of large wall-thickness differences.
- Close-joint soldering: radiators, pipes in the assembly area.

Manufacturing methods

The method used for heating provides yet another criterion for describing soldered joints. The most important types are furnace, induction, open-flame, and iron soldering.

Sweating

Heating is performed in through-type or vacuum furnaces with defined temperature and time characteristics. Before entering the furnace, the components are secured and the solder is inserted or applied as a paste. Additional fluxes are not required. The furnace atmosphere and the vacuum assume the flux function without leaving residues on the components.

Induction soldering

Heating is locally limited by inductive means.

Induction soldering is mainly used for round components, e.g., pipes for cooling water or oil.

Flame soldering

Heating is performed by individual torches or in a gas-heated system. Depending on the specific soldering operation, oxyacetylene burners (familiar from gas welding), propane torches or soldering lamps are employed. The solder is usually supplied from the side in the shape of a rod. The flux is applied separately or integrated in the solder.

Flame soldering is used for very small production batches and prototypes or to rework and repair, for example, pipes.

Iron soldering

A hand-held or mechanically guided soldering iron provides the heat. Irons can also be used to solder pretinned surfaces.

Iron soldering only used in soft soldering.

Further processes

- Salt-bath soldering,
- Dip soldering,
- Resistance soldering, as well as,
- MIG soldering,
- Plasma soldering, and
- Laser soldering.

Close-joint soldering and**V-joint brazing**

Furthermore, with regard to the geometry of the solder and the soldering flux, a distinction is made between close-joint soldering and V-joint brazing.

Close-joint soldering

The soldering gap is filled by the capillary effect of the liquid solder. The strength of the soldered joint is determined entirely by the strength of the base materials. Typical manufacturing processes are sweating, flame soldering and induction soldering.

V-joint brazing

The soldering gap is filled by the force of gravity. The strength of the soldered joint is determined above all by the strength of the solder filler. Typical manufacturing processes are arc and laser soldering.

Adhesive technologies**Adhesives**

An adhesive is a non-metallic material which can join workpieces (parts to be joined) by means of adhesion and its internal strength (cohesion) without the structure of the part to be joined being significantly altered. An adhesive enables the creation of permanent material joints between homogeneous or heterogeneous material pairings.

Organic and inorganic adhesives as they harden build up firstly adhesion with the material surfaces by way of chemico-physical interactions (adhesion or adhesive forces) and secondly the necessary structural strength (cohesion). Depending on their chemistry, these adhesives harden for example at room temperature, at increased temperatures or under UV radiation with different mechanisms.

Hardening (cross-linking) and adhesion occur, in accordance with the chemical basic structure and formulation, over a certain period of time, depending on for example temperature or air humidity. The reaction types are called polymerization, polyaddition and polycondensation. Spatially interlinked macromolecules are the result. Adhesives can, depending on the required hardening temperature, be subdivided into the category of those that harden at room temperature or at higher temperatures (100 to 200 °C).

Two different types of adhesive are used, based on their recipe and form of presentation: single-component and two-component adhesives.

Two-component adhesives

These are adhesives which are made up of two components which must be mixed together to harden. The stoichiometric mixture ratio of the two components must be maintained exactly. Component A contains the basic resin. The second component (component B) contains a hardener which initiates the cross-linking reaction and links the resin molecules from component A to each other. An accelerator may be added to the hardener. Hardening usually occurs at room tem-

perature, but can be accelerated by the additional influence of slightly increased temperatures (80 to 120 °C) and thus completed more quickly.

Single-component adhesives

Single-component adhesives contain all the constituents necessary to form a bond in one phase. These systems contain inhibitors which prevent a premature chemical reaction (hardening) between the reacting agents present in the one phase (monomers, resin and hardener; premixed two-component adhesive). Single-component adhesives do not require the additional process step of mixing the components as a two-component adhesive requires.

Hardening must, depending on the chemical formulation, be initiated by higher temperatures (in an oven, by induction or infrared radiation), ultraviolet radiation or air humidity, and completed. In this process the inhibitors are neutralized and the accelerators also contained in the adhesive are released to speed up the hardening process.

Most single-component adhesives must be stored under cooled conditions (electrical conductive adhesives down to -20 °C in an upright deep-freezer, structural adhesives and sealants at 4 to 10 °C in a cold warehouse) or in a dark environment (UV adhesives) until they are ready for use.

Structural designs of adhesive connections

Adhesive connections should be designed so that predominantly shear loads occur. Overlapping joints are particularly suitable. Butt joints subjected to tensile or sliding forces should be avoided.

Combinations of adhesive bonding with other joining methods, e.g. welding, screwing or riveting, can have beneficial effects. The joining spots secure the components during the adhesive hardening time. Stress concentrations for example at the edges of weld spots or at rivet spots can also be minimized. Constructions of this nature provide enhanced structural integrity, rigidity, and damping when subjected to dynamic loads.

Examples of adhesives with technical application potential

The most important adhesives are epoxy resins, silicones, polyurethanes, and acrylates.

Silicone and polyurethane adhesives

Silicone and polyurethane adhesives are important particularly in bonds with dynamic loads, but also where tightness against liquid media (e.g. water) is called for. Depending on the medium (polar, nonpolar), there is however a risk of flexible adhesives swelling (e.g. in fuel). This process can be reversed provided the system can be re-dried during operation and the adhesive has not suffered any damage (e.g. material embrittlement by extraction). Thanks to their flexibility (elastic properties), these adhesives can compensate the different coefficients of thermal expansion of the parts to be joined in the temperature-application range by means of their deformability (low modulus of elasticity above a crystallization temperature of -60 °C to -40 °C) such that only low mechanical stresses can be built up in the bond. Compared with polyurethanes, silicones are also highly suitable for use at high application temperatures up to approximately 220 °C.

Epoxy adhesives

Epoxy adhesives, on the other hand, are predominantly very brittle and hard adhesives which can, depending on their formulation, be used up to approximately 200 °C. Compared with flexible adhesives, they are much less susceptible to swelling in liquid media. High-strength epoxy resins lack the ability of silicones to compensate by means of deformability (elongation) mechanical stresses in the adhesive bond which are caused by different coefficients of thermal expansion of the parts to be joined in the temperature-application range. Exceptions are adhesives which, thanks to their special formulation (e.g. impact-resistant modification), are used specifically for bonds in body manufacturing in order specifically to reduce the mechanical forces and energies that occur in a crash situation and to maintain the structural integrity of the passenger cell. Epoxy adhesives facilitate bonds with

very high adhesive strengths (high modulus of elasticity up to the glass-transition temperature; this is in the range of 80 to 200 °C, depending on the chemical recipe of the adhesive components).

Acrylate adhesives

Acrylates are capable of fast hardening reactions and therefore highly attractive for use in production processes. These adhesives can harden, depending on their chemical architecture, in response to ultraviolet radiation, mixing of components, at increased temperatures or also in response to air humidity ("superglue") within a few minutes down to a matter of seconds. On the other hand, acrylates do not demonstrate as pronounced thermomechanical performance as silicone and epoxy adhesives above approximately 120 °C or in aggressive media.

Automotive applications

Adhesive joining has become a standard technique in automotive engineering. The individual areas of application can be classified as follows:

- Electronic components: sealing/bonding of housings for sensors, ECUs and video systems; electrically and thermally conductive bonding of electronic component in power modules and electronic circuits.
- Electric motors: magnetic bonds in rotors and stators.
- Body shell: raised-seam and brace bonding for attached components.
- Assembly line: attachment of insulating material, appliqué, moldings, mirror-support bracket to windshield.
- Component production: bonding of brake pads, laminated safety glass (LSG), rubber-metal connections to absorb vibration.

Riveting

Classic riveting

Process

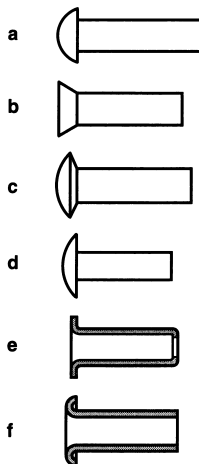
Riveting is used to produce a permanent fixed connection between two or more components made of identical or dissimilar materials. Here the components to be joined are pierced together by means of drilling or punching. The rivet is then inserted into the hole as the joining element. Depending on the method and application, riveted joints are divided into the following categories:

- Permanent rigid connections (interference fits, for example in mechanical and plant engineering),
- Permanent, sealed connections (for example in boilers and pressure vessels),
- Extremely tight seals (for example in pipes, vacuum equipment).

A distinction is drawn between cold and hot riveting, depending on the temperature used. Cold riveting is employed for rivet joints up to 10 mm in diameter in steel, copper, copper alloys, aluminum,

Figure 5: Rivet shapes

- a) Mushroom-head rivet,
- b) Countersunk-head rivet,
- c) Oval countersunk-head rivet,
- d) Flat round-head rivet,
- e) Hollow rivet,
- f) Tubular rivet.



etc. Rivets with diameters in excess of 10 mm are installed hot.

The most common rivet shapes are (Figure 5, [15]...[20]) the mushroom-head rivet (DIN 660), the countersunk-head rivet (DIN 661), the oval countersunk-head rivet (DIN 662), the flat round-head rivet (DIN 674), the hollow rivet (DIN 7339), and the tubular rivet (DIN 7340).

There are also standardized rivets for specialized applications, e.g. explosive rivets or blind rivets. Blind rivets are hollow and are expanded by a drift or punch.

Furthermore, rivets are frequently designed for use as function elements. Examples are rivet nuts and clinch bolts, which serve as bolting points.

The strength properties and the chemical composition of rivet materials are laid down in numerous national and international standards. In the interests of avoiding electrochemical corrosion, it is advisable wherever possible to use materials of the same kind for rivets and component.

In mechanical engineering in general and in tank manufacture riveting has largely been displaced by welding.

Advantages and disadvantages compared with other joining techniques

- Unlike welding, riveting exerts no influence such as hardening or structural transformation on the material.
- No distortion of the components.
- Suitable for joining dissimilar materials.
- Riveting weakens the components.
- Butt joints cannot be riveted, and
- Riveting is generally more expensive than welding when performed outside the factory.

Automotive applications

- Riveting pivot/joint pins (power-window units, hinges, windshield-wiper linkages).
- Riveting reinforcement plates (in the course of repairs).

Punch riveting

Process

The punch-riveting technique joins solid materials employing stamping and riveting elements (solid or semi-hollow rivets) without piercing in a combined cutting and joining operation. The parts to be joined do not have to be pierced or predrilled, as is the case with other riveting techniques.

Punch riveting with semi-hollow rivets

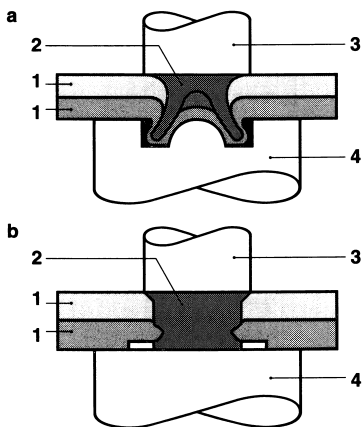
The first stage in punch riveting with a semi-hollow rivet (Figure 6a) is to position the joining point of the components to be joined on the (bottom) die plate. The riveting die descends and presses the semi-hollow rivet through the upper sheet into the lower sheet in a single stamping operation. The rivet deforms and the bottom spreads to form a securing element, usually without fully penetrating the lower sheet.

Punch riveting with solid rivets

The first stage in punch riveting with a solid rivet (Figure 6b) is to position the joining point of the components to be joined on the (bottom) die plate. The top section of the rivet unit, including the blank holder, descends and the riveting die presses the rivet through the parts to be joined in a single stamping operation.

Figure 6: Joint with punch rivet

- a) Semi-hollow rivet,
b) Solid rivet.
1 Workpieces,
2 Punch rivet,
3 Riveting die,
4 Die plate.



Equipment

Hydraulic joining equipment is used to produce the joints; in equipment of this type the riveting die and the die plate are arranged in a very rigid, C-shaped frame. The rivets can be supplied to the riveting tool loose or via magazined carrier strips.

Materials

The rivets must be harder than the materials to be joined. The most common materials are steel, stainless steel, copper and aluminum with various coatings.

Features

- Possible to join similar and dissimilar materials (e.g. steel, plastic or aluminum), parts of various thicknesses and strengths, and painted sheets.
- No preliminary piercing or drilling, no heat or vacuum extraction required.
- The total material thickness that can be riveted is 6.5 mm for steel and 11 mm for aluminum.
- The joining process generates minimal heat and noise.
- Tools have a long service life (approx. 300,000 rivet applications) and deliver consistent joint quality over a long period of time.
- High process reliability through monitoring of the process parameters.
- High forces have to be applied.
- Larger tong projections of the riveting tools are only possible under limited conditions owing to the rigidity requirements.

Applications

- Punch riveting with solid rivets: joining metal sheets in automotive engineering, e.g. power-window drives in passenger cars.
- Punch riveting with semi-hollow rivets: joining materials in passenger-car body manufacturing, white goods (household appliances), joining metals to composites (heat shields).

Penetration-clinching processes

Processes

Penetration clinching ("clinching") comprises mechanical joining processes which combine penetration, cold upsetting and sometimes also cutting in a single continuous joining operation without the application of additional heat. Based on this principle, it can be assigned to the process of joining by forming (see DIN 8593-5, [21]).

Distinctions are made between processes with and without cutting and those where the joining point is round or rectangular in shape.

"Tox clinching"

Some process variations are still referred to in technical parlance by the original manufacturer's designations, e.g. "Tox clinching" for penetration clinching with a round die without cutting (Figure 7b). The tools employed for "Tox clinching" are relatively small. The diameter can be varied to suit specific applications. A force-travel curve typical of "Tox clinching" can be broken down into five characteristic phases (A-E) (Figure 8).

Penetration clinching

Penetration clinching (Figure 7a) currently allows blanks up to 3 mm thick to be joined, whereby the total thickness of the two blanks should not exceed 5 mm. The blanks being joined can be of the same material (e.g. steel to steel) or dissimilar materials (e.g. steel to nonferrous metal). Penetration clinching can be used to join coated sheets and painted parts, as well as components to which adhesives have been applied. Multiple penetration clinching can be used to produce numerous elements to be joined (up to 50) in a single process (one stroke of the press, for example).

Advantages and disadvantages of penetration-clinching processes

- No need for an acoustic enclosure.
- Tox clinching to a large extent does not impair corrosion protection.
- When combined with cutting, there is a partial loss of corrosion protection.
- No distortion due to thermal stress.
- Painted, protected (oil, wax) and glued sheets can be clinched.
- Different materials can be joined (e.g. steel to plastic).
- Energy savings, since no power supply for welding and no cooling-water system are required.

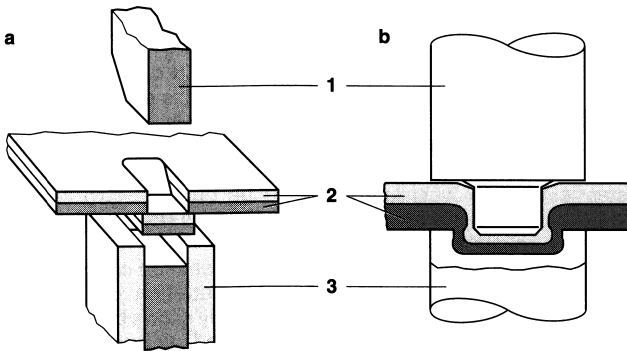
- One side of the workpiece has a projection similar to that produced by a rivet head, while the other side has a corresponding depression.

Automotive applications

- Steel and aluminum bodies,
- Windshield-wiper brackets,
- Fastening interior door panels,
- Hinges, locks,
- Seat systems.

Figure 7: Penetration-clinching processes

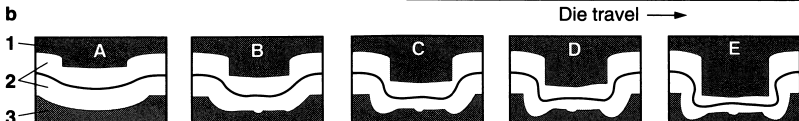
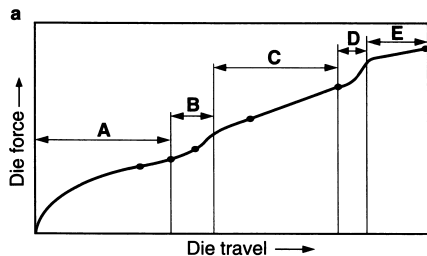
- a) Clinching,
b) "Tox clinching".
1 Die,
2 Parts to be joined,
3 Die plate.



UAM00592Y

Figure 8: Typical force/travel curve in penetration-clinching processes

- a) Die-force/die-travel curve,
b) Process steps.
A Combined indentation and penetration,
B Upsetting and expansion,
C Filling upper contour of mold,
D Filling annular groove of mold,
E Cup extrusion.
1 Die,
2 Parts to be joined,
3 Die plate.



UAM0059-3E

References

- [1] Fügetechnik Schweißtechnik; 7th Edition, DVS Media, 2007.
- [2] DIN-DVS Taschenbuch 284 – Schweißtechnik 7: Schweißtechnische Fertigung, Schweißverbindungen; 3rd Edition, DVS Media, 2009.
- [3] DIN 1910-100:2008: Welding and allied processes – Vocabulary – Part 100: Metal welding processes with additions to DIN EN 14610:2005.
- [4] DIN EN 14610:2005: Welding and allied processes – Definitions of metal welding processes; trilingual version EN 14610:2004.
- [5] DIN-DVS Taschenbuch 283 – Schweißtechnik 6: Strahlschweißen, Bolzenschweißen, Reibschweißen; 4th Edition. DVS Media, 2009.
- [6] DIN EN 1011: Welding – Recommendations for welding of metallic materials – Part 6 (2006): Laser beam welding; German version EN 1011-6:2005. Part 7 (2004): Electron beam welding; German version EN 1011-7:2004.
- [7] DIN-DVS Taschenbuch 196/1 – Schweißtechnik 5: Hartlöten; 5th Edition. DVS Media, 2008.
- [8] DIN-DVS Taschenbuch 196/2 – Schweißtechnik 12: Weichlöten, gedruckte Schaltungen; 1st Edition. DVS Media, 2008.
- [9] DIN ISO 857-2:2007: Welding and allied processes – Vocabulary – Part 2: Soldering and brazing processes and related terms (ISO 857-2:2005).
- [10] DIN EN ISO 9454-1:2016: Soft soldering fluxes; classification and requirements; part 1: classification, labeling and packaging (ISO 9454-1:2016); German version EN ISO 9454-1:2016.
- [11] DIN EN 1045:1997: Brazing – Fluxes for brazing – Classification and technical delivery conditions; German version EN 1045:1997.
- [12] DIN EN ISO 9453:2014: Soft solder alloys – Chemical compositions and forms (ISO 9453:2014); German version EN ISO 9453:2014.
- [13] DIN EN ISO 12224-1:1998: Solid wire, solid and flux cored – Specification and test methods – Part 1: Classification and requirements (ISO 12224-1:1997); German version EN ISO 12224-1:1998.
- [14] DIN EN ISO 17672:2010: Brazing – Filler metals (ISO 17672:2010); German version EN ISO 17672:2010.
- [15] DIN 660:2012: Round head rivets – Nominal diameters 1 mm to 8 mm.
- [16] DIN 661:2011: Countersunk head rivets – Nominal diameters 1 mm to 8 mm.
- [17] DIN 662:2011: Mushroom head rivets – Nominal diameters 1.6 mm to 6 mm.
- [18] DIN 674:2011: Flat round head rivets – Nominal diameters 1.4 mm to 6 mm.
- [19] DIN 7339:2011: Hollow rivets, one piece, draw from strip.
- [20] DIN 7340:2011: Tubular rivets cut from the tube.
- [21] DIN 8593-5:2003: Manufacturing processes joining – Part 5: Joining by forming processes; Classification, subdivision, terms and definitions.